

Федеральное государственное автономное
образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»
Институт космических и информационных технологий
Кафедра «Информатика»
НУЛ «Системный анализ и управление»

УТВЕРЖДАЮ
Заведующий кафедрой
_____ А.С. Кузнецов _____
подпись инициалы, фамилия
«__» _____ 2017 г.

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Оптимизация сети передачи данных на примере Сибирского банка ПАО
Сбербанка России
27.04.03 Системный анализ и управление
27.04.03.02 Системный анализ данных и технологий принятия решений

Научный руководитель	_____	ДОЦЕНТ, К.Т.Н.	А.А. Даничев
	подпись, дата	должность, ученая степень	инициалы, фамилия
Выпускник	_____		М.Д. Тюрюмина
	подпись, дата		инициалы, фамилия
Рецензент	_____	вед. инж. ЦИ ПАО Сбербанк	Д.О. Видинеев
	подпись, дата	должность, ученая степень	инициалы, фамилия

Красноярск 2017

РЕФЕРАТ

Выпускная квалификационная работа по теме «Оптимизация сети передачи данных на примере Сибирского банка ПАО Сбербанк России» содержит 46 страниц текстового документа, 34 использованных источника, 13 иллюстраций, 11 формул.

МНОГОКРИТЕРИАЛЬНАЯ ОПТИМИЗАЦИЯ, МОДЕЛЬ, СЕТИ ПЕРЕДАЧИ ДАННЫХ, КОНЦЕПЦИЯ ПАРЕТО, МЕТОД ВЕТВЕЙ И ГРАНИЦ, МЕТОД ИМИТАЦИИ ОТЖИГА, ПРОГРАММА.

Объектом изучения является сеть передачи данных и методы ее оптимизации.

Целью является разработка и реализация алгоритма, способного находить с приемлемой точностью решение задачи многокритериальной оптимизации сети передачи данных Сибирского банка при большой размерности.

В результате данной работы была построена модель сети передачи данных, поставлена задача бикритериальной оптимизации, выполнен обзор существующих методов поставленной задачи, выбраны и реализованы наиболее подходящие методы, на основе результатов исследования работы алгоритмов выбран наиболее подходящий из них, предложены модификации в связи со спецификой предметной области.

В итоге разработан модифицированный алгоритм, позволяющий находить множество оптимальных в определенном смысле решений модернизации сети передачи данных Сибирского банка за приемлемое время.

СОДЕРЖАНИЕ

Введение.....	6
1 Методы решения многокритериальных задач оптимизации сети передачи данных	8
1.1 Многокритериальная оптимизация	8
1.2 Подходы к решению	9
1.2.1 Свертывание критериев.....	9
1.2.2 Концепция Парето	10
1.3 Метод ветвей и границ в решении задач оптимизации.....	11
1.4 Эвристические алгоритмы	12
1.5 Задачи транспортного типа	13
1.6 Моделирование сети передачи данных. Методы оптимизации	14
1.7 Выводы по главе 1.....	15
2 Математическая модель, задача и алгоритмы оптимизации сети передачи данных	15
2.1 Структура сети передачи данных банка	15
2.2 Математическая модель сети передачи данных	16
2.3 Задача оптимизации сети передачи данных	17
2.4 Определение множества Парето-оптимальных решений задачи оптимизации сети передачи данных путем полного перебора.....	19
2.5 Определение множества Парето-оптимальных решений задачи оптимизации сети передачи данных методом ветвей и границ	19
2.5.1 Пример решения задачи методом ветвей и границ	23
2.6 Выводы по главе 2.....	27

3 Поиск субоптимального множества решений задачи оптимизации сети передачи данных	27
3.1 Субоптимальное множество	28
3.2 Оценка отклонения субоптимального множества решений от множества Парето.....	28
3.2.1 Метод подсчета решений	29
3.2.2 Метод оценки точности приближения субоптимального множества решений к Парето-оптимальному	30
3.3 Поиск субоптимального множества решений с применением имитации отжига.....	31
3.3.1 Начальный вектор решений. Получение нового решения	32
3.3.2 Снижение температуры. Вычисление энергии.	32
3.3.3 Критерии допуска и критерий останова	33
3.3.4 Дополнительные эвристики	34
3.3.5 Оценка вычислительной сложности	34
3.4 Комбинированный метод	34
3.5 Вычислительный эксперимент	35
4 Программный комплекс решения задачи проектирования и оптимизации сети передачи данных на примере Сибирского банка.....	39
4.1 Назначения и возможности программного комплекса	Error! Bookmark not defined.
4.3 Системные требования	39
4.4 Описание графического пользовательского интерфейса.....	40
4.4.1 Главное окно программы	40
4.4.2 Главное меню программы	42
4.5 Окно ввода данных о канале связи.....	44

4.6 Типовой сценарий работы с программным комплексом	45
4.7 Рекомендации по использованию программного комплекса	46
Список использованных источников	51
Приложение А	55
Приложение Б	56

ВВЕДЕНИЕ

С каждым годом объемы передаваемой информации посредством сети Интернет стремительно растут. Соответственно возрастает нагрузка на сети передачи данных (далее СПД), что служит толчком для усовершенствования и оптимизации.

Интернет-провайдеры (далее провайдеры) стремятся быть конкурентоспособными, поэтому качество предоставляемых услуг улучшается, изменяется стоимость аренды каналов, появляются новые каналы связи и т.д.

Иногда изменения в тарифных планах существенны и выгодны для пользователей, но проанализировать и решить, стоит ли воспользоваться теми или иными услугами довольно проблематично.

Рассмотрим Сибирский банк - филиал ПАО Сбербанк России. Он обслуживает клиентов на территории Сибирского федерального округа.

С целью повышения качества сервиса и эффективности обслуживания банк активно развивает как существующие технологии предоставления услуг, так и внедряет новые. Например, услуга приема платежей у населения реализована и в банкоматах, и в информационно-платежных терминалах Банка. Связь между банкоматами, терминалами, отделениями и центром обработки данных обеспечивается с помощью арендованных у провайдера каналов связи.

Проанализировав деятельность отдела систем связи и телекоммуникаций было выявлено, что существует нехватка программного обеспечения, готового анализировать предложения провайдеров и предоставлять альтернативные решения при определении плана аренды дополнительных каналов для улучшения технических характеристик СПД.

Банк заинтересован в получении качественных услуг от провайдеров за минимально возможную стоимость. В этом случае необходимо учитывать два различных критерия. Таким образом, можно определить задачу бикритериальной оптимизации.

Объектом исследования является сеть передачи данных и методы её оптимизации.

Предметом исследования является математическая модель и алгоритмы оптимизации СПД Сибирского банка.

Целью работы является разработка и реализация алгоритма, способного находить с приемлемой точностью решение задачи многокритериальной оптимизации сети передачи данных Сибирского банка при большой размерности.

Для достижения поставленной цели решались следующие задачи:

- моделирование существующей СПД Сибирского банка;
- определение основных критериев оптимальности, постановка задачи бикритериальной оптимизации сети;
- анализ существующих методов решения многокритериальных задач и выбор наиболее подходящих для решения рассматриваемой практической задачи;
- программная реализация выбранных методов;
- разработка и реализация программы с графическим интерфейсом для нужд Сибирского банка.

1 Методы решения многокритериальных задач оптимизации сети передачи данных

1.1 Многокритериальная оптимизация

Многокритериальная оптимизация (далее МКО) – это процесс одновременной оптимизации двух и более конфликтующих целевых функций в заданной области определения.

Таким образом, задачу выбора оптимальных характеристик СПД можно считать многокритериальной, поскольку необходимо рассматривать как минимум два критерия: пропускную способность сети и арендную стоимость.

МКО представляет собой попытку получить наилучшее значение для некоторого множества характеристик рассматриваемого объекта, найти некий компромисс между теми частными критериями $Q_i(\vec{x})$ (где $i = 1, 2, \dots, s$), по которым требуется оптимизировать решение.

В условиях определенности, то есть в отсутствие случайных или неопределенных событий, задачу МКО можно представить следующим образом:

$$\min_{x \in D} (f_1(x)), \dots, \min_{x \in D} (f_n(x)), \quad (1)$$

где D – множество допустимых решений;

$f_1(x), \dots, f_n(x)$ ($n \geq 2, x \in D$) – множество числовых функций, определенных на множестве D .

Числовые функции задают векторный критерий $\vec{f} = (f_1(x), \dots, f_n(x))$, принимающий значения в пространстве n -мерных векторов \mathbb{R}^n , где \mathbb{R} – множество вещественных чисел.

Оценкой решения $x \in D$ называют всякое значение $f(x) = (f_1(x), \dots, f_n(x))$. В задаче МКО сравнение решений происходит с помощью векторного критерия f .

Далее приведены наиболее популярные подходы к решению задач МКО.

1.2 Подходы к решению

1.2.1 Свертывание критериев

Широко известным является подход к решению многокритериальных задач, основанный на «свертывании» векторного критерия f в одну функцию – обобщенный критерий $F(f_1, \dots, f_m)$. Это значительно упрощает решение задачи, поскольку происходит сужение множества решений от \mathbb{R}^n до \mathbb{R} (n – число критериев). Наиболее распространенным является метод линейной комбинации частных критериев.

Пусть задан вектор весовых коэффициентов критериев $\alpha = (\alpha_1, \dots, \alpha_k)$, характеризующих важность соответствующего критерия, $\sum_{k=1}^K \alpha_k = 1, \alpha_k \geq 0, k = \overline{1, K}$. Линейная скаляризованная функция представляет собой сумму частных критериев, умноженных на весовые коэффициенты. Задача математического программирования становится однокритериальной и имеет вид

$$F^0 = \sum_{k=1}^K \alpha_k f_k(x) \rightarrow \max, g_i(x) \leq b_i, i = \overline{1, I}, x \geq 0. \quad (2)$$

Критерии в свертке могут быть нормированы. Решение, полученное в результате оптимизации скаляризованного критерия эффективно.

К недостаткам метода можно отнести то, что малым приращениям коэффициентов соответствуют большие приращения функции, т. е. решение задачи неустойчиво. Кроме того, необходимо определение весовых коэффициентов.

1.2.2 Концепция Парето

Данный подход позволяет находить не одно, а целое множество оптимальных в определенном смысле решений. Рассмотрим оценки $x' = (f_1(x'), \dots, f_n(x'))$ и $x'' = (f_1(x''), \dots, f_n(x''))$ задачи (1). Если $f_i(x') \leq f_i(x'')$, $i = 1, 2, \dots, n$, и хотя бы одно неравенство выполняется строго, то говорят, что оценка $f(x')$ доминирует $f(x'')$ и обозначают $f(x') \succ f(x'')$. Решение x' доминирует решение x'' , если соответствующая ему оценка доминирует оценку решения x'' . Такое отношение обозначается как $x' \succ x''$.

Если множество решений D не содержит такого x'' , что $f(x'') \succ f(x')$, то оценка $f(x')$ называется Парето-оптимальной (недоминируемой). Совокупность всех Парето-оптимальных решений X задачи оптимизации есть множество Парето.

Решение задачи с применением концепции Парето сводится к следующим шагам:

- 1) определение множества Парето-оптимальных решений;
- 2) выбор оптимальной оценки с точки зрения лица, принимающего решение (ЛПР);
- 3) построение решения задачи, соответствующее выбранной оценке.

Ввиду того, что ЛПР имеет определенный опыт в своей предметной области, возможность выбора из множества оптимальных по Парето решений позволяет ему применить дополнительные неформализованные знания и получить лучший, по его мнению, результат.

Для определения множества Парето предлагается использовать адаптированные под решение многокритериальных задач метод ветвей и границ [14, 33, 34, 35], эвристические подходы на основе алгоритмов имитации отжига, генетических алгоритмов, нейронных сетей и другие.

Свойства и методы решения многокритериальных задач более подробно описаны в работах [3, 7, 9, 23, 26, 27, 35].

1.3 Метод ветвей и границ в решении задач оптимизации

В 1960 году А. Лэнд и Дж. Дойг предложили новый метод решения задач линейного программирования – метод ветвей и границ.

Метод ветвей и границ – это, своего рода, вариант метода полного перебора с отбрасыванием заведомо неоптимальных подмножеств допустимых решений.

В общем случае метод состоит из следующего набора действий:

1) определение множества D допустимых решений задачи и формирование списка ветвления.

2) непосредственное выполнение процесса ветвления. Разбиение множества D на произвольное число k подмножеств D_i ($\bigcup_{i=1}^k D_i = D$; $\bigcap_{i=1}^k D_i = \emptyset$) с последующим добавлением их в список кандидатов для рассмотрения. Пустой список кандидатов говорит о том, что все решения были рассмотрены или отсеяны. В таком случае алгоритм завершает свою работу.

3) выбор подмножества для ветвления из полученного списка кандидатов. Вычисление «верхней» и «нижней» оценки. В случае, когда «верхняя» оценка больше (меньше) значения рекорда при рассмотрении задачи максимизации (минимизации), происходит разбиение рассматриваемого подмножества аналогично шагу 2 с последующим обновлением списка кандидатов на ветвление. В противном случае, ветвление для данного подмножества завершается.

4) обновление значения рекорда по необходимости.

5) отсеивание множеств, не содержащих оптимальные решения, и проверка признака оптимальности. Переход к шагу 3.

Также возможно совместное применение нескольких подходов, например, метода ветвей и границ и концепции динамического программирования. Такие подходы называют комбинированными.

1.4 Эвристические алгоритмы

Эвристическими называют такие алгоритмы решения задачи, правильность которых для всех возможных случаев не доказана, но известно, что большинство полученных решений достаточно хорошие. Может быть даже известно (то есть доказано), что эвристический алгоритм формально неверен. Его все равно можно применять, если результат только в отдельных, достаточно редких и хорошо выделяемых случаях неверен или же неточен, но все же приемлем.

Другими словами, это не полностью математически обоснованные (или даже «не совсем корректные»), но при этом практически полезные алгоритмы.

Следует обратить внимание на то, что эвристический алгоритм, в отличие от корректного алгоритма решения задачи, обладает следующими особенностями:

- не гарантирует нахождение лучшего решения;
- не гарантирует нахождение решения, даже если оно заведомо существует (возможен «пропуск цели»);
- может дать неверное решение в некоторых случаях.

Эвристические алгоритмы широко применяются для решения задач высокой вычислительной сложности (задачи, принадлежащие классу NP), то есть вместо полного перебора вариантов, занимающего существенное время, а иногда технически невозможного, применяется значительно более быстрый, но недостаточно обоснованный теоретически алгоритм.

Однако эвристические алгоритмы имеют преимущества перед другими алгоритмами при очень больших размерностях задач и отсутствии упорядоченности в исходных данных, когда альтернативой им является метод полного перебора вариантов. Главным преимуществом является то, что они могут применяться для решения сложных неформализованных задач, для которых не разработано специальных методов, т.е. обеспечивается решение нестандартных проблем.

К числу эвристических методов оптимизации относят алгоритм имитации отжига, генетические алгоритмы, эволюционные и поведенческие стратегии и так далее.

1.5 Задачи транспортного типа

Задачи транспортного типа составляют специальный и довольно обширный класс задач линейного программирования.

Возникающая в различных предметных областях проблема заключается в необходимости выбора оптимального способа транспортировки некоторого продукта из одного или нескольких источников множеству стоков.

Впервые проблема была формализована французским математиком Гаспаром Монжем в 1781 году. В 1920-х А.Н. Толстой был одним из первых, кто занимался ее решением. Формированию класса задач транспортного типа способствовали работы Ф.Л. Хичкока, Л.В. Канторовича [11] и М.К. Гавурина. Начиная с работ Дж. Данцига ведутся исследования по определению максимальной пропускной способности (максимального потока), потока минимальной стоимости в сети и решению других задач рассматриваемой области.

Немало работ посвящены исследованию задач оптимизации телекоммуникационных сетей [6, 10, 37].

Классические модели теории оптимизации телекоммуникационных сетей используют матрицу потребностей (или нагрузки), которая содержит данные о количестве информации, передаваемой между узлами сети, и некоторую функцию, зависящую от величины задержки пакета, в качестве критерия оптимизации.

В данной работе рассматривается модель, отличающаяся некоторым образом от классических.

Во-первых, в модели не используется матрица потребностей. Это обусловлено появлением безлимитных тарифов с фиксированной абонентской

платой. В этом случае требования, задаваемые матрицей потребности, больше не актуальны.

Во-вторых, вместо одного из критериев оптимальности – функции задержки пакета – используется максимальная пропускная способность сети. За последний десяток лет пропускная способность каналов передачи данных увеличилась в несколько раз, а размеры сетевых пакетов остались прежними. Таким образом, целесообразно использовать концепцию потоков для моделирования передачи сетевых пакетов.

В-третьих, эффективность проектировочного решения оценивается по значению не одного, а двух критериев и ищется не один, а множество альтернативных вариантов. Это отличие обусловлено тем, что в регламенте Сибирского банка прописаны процедуры увязки технических решений с затратами на их реализацию и необходимости множества альтернативных вариантов при выборе плана модернизации.

1.6 Моделирование сети передачи данных. Методы оптимизации

На сегодняшний день в моделировании СПД широко применяются элементы теории графов, понятие потока в сетях и системы с очередями [12].

Топология сети описывается графом $G = (V, E)$, где множество вершин $V = \{v_1, \dots, v_n\}$ – это узлы коммутации сети, множество ребер $E = \{e_1, \dots, e_m\}$ – каналы передачи данных, характеризующиеся пропускной способностью c_k ($k = 1, 2, \dots, m$) и интенсивностью потоков f_k ($k = 1, 2, \dots, m$), γ – суммарный трафик по сети [12]. Предлагается оценка средней задержки сообщений на каналах при передаче:

$$T = \frac{1}{\gamma} \sum_{k=1}^m \frac{f_k}{c_k - f_k}. \quad (3)$$

Формулировка задачи проектирования сети звучит следующим образом: заданы множества коммуникационных узлов $V = \{v_1, \dots, v_n\}$ и матрица внешней нагрузки $\Gamma = (\gamma_{i,j})$, которая обеспечивает заданную связность графа $G = (V, E)$ и минимизирует суммарную стоимость сети при ограничении средней задержки на каналах связи.

Для решения данной задачи существует ряд методов: вогнутый метод удаления ребер (ВМУР) [10, 12], метод групповой оптимизации и итеративный алгоритм М-структур [10] и другие.

1.7 Выводы по главе 1

В главе были даны общие сведения о задачах МКО, задачах транспортного типа, а также проведен обзор некоторых подходов к их решению. Рассмотрена общая модель и некоторые методы оптимизации СПД.

2 Математическая модель, задача и алгоритмы оптимизации сети передачи данных

2.1 Структура сети передачи данных банка

Основными поставщиками телекоммуникационных услуг для Сибирского банка, работающими на территории Сибирского федерального округа являются «Ростелеком», «Мегафон» и «ТрансТелеКом». Банк арендует высокоскоростные каналы связи по безлимитному тарифу.

Типовая схема подключения отделений банка, терминалов и банкоматов к центру обработки данных (ЦОД) представлена на рисунке 1. В качестве примера на рисунке В.1 представлена географическая карта части Республики Хакасии с нанесенными на нее узлами коммутации, арендованными банком каналами и каналами связи, предлагаемые провайдером «Ростелеком» к аренде в 2015 году.

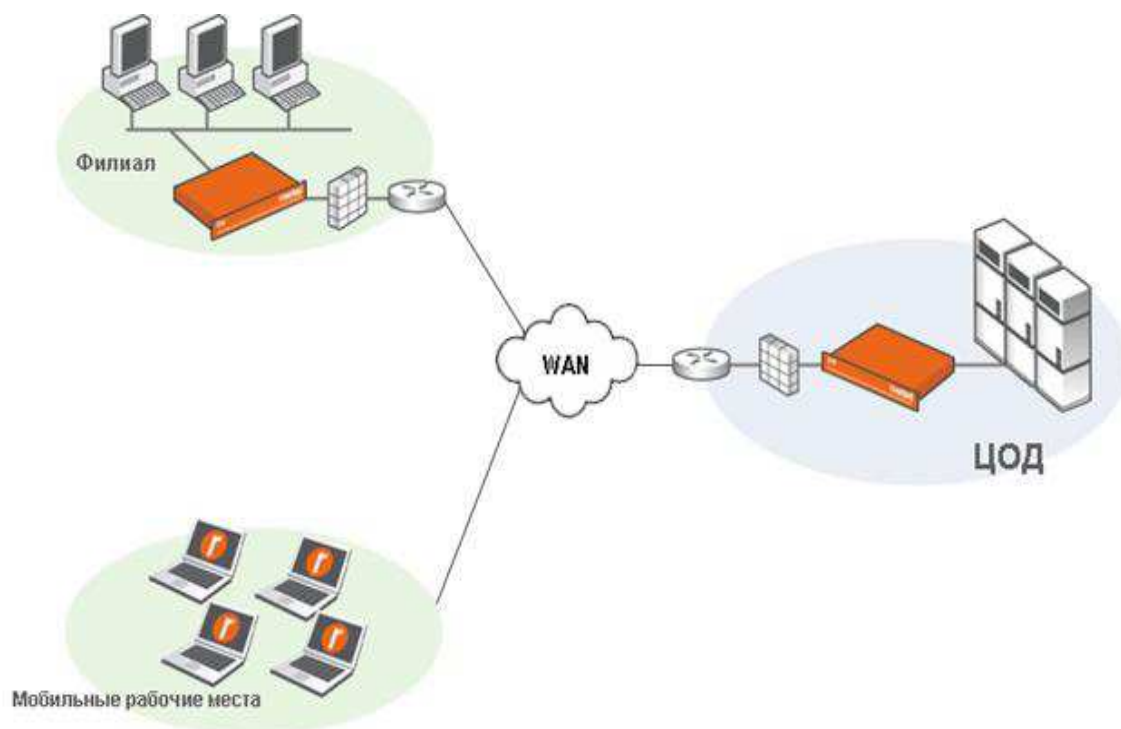


Рисунок 1 – Типовая схема подключения

2.2 Математическая модель сети передачи данных

Процедура оптимизации базируется на модели СПД, основанной на классических потоковых моделях [15, 34].

Существующую СПД задает ориентированный ациклический граф $G = (V, E)$ (V и E - множества вершин и ребер соответственно). Вершинам соответствуют узловые элементы сети (маршрутизаторы). Каждое ребро $(u, v) \in E$ характеризуется положительной пропускной способностью $c(u, v)$, обусловленной технической спецификацией конкретного канала передачи данных. В графе выделяются источник информации s и сток t . Они могут быть как реальные, так и фиктивные.

Для описания каналов передачи данных, которые предлагает провайдер к аренде, задается множество ребер E' ($E \cap E' = \emptyset$). Для каждого $(u, v) \in E'$ известна пропускная способность $c'(u, v)$ и стоимость аренды $p'(u, v)$.

Функция потока $f_G: V \times V \rightarrow \mathbb{R}$ представляет собой модель процесса передачи данных в сети. Величина $f_G(u, v)$ показывает количество информации, переданное в единицу времени по каналу связи между узлами u и v . Функция f_G должна удовлетворять условиям:

- ограниченность пропускной способности: $\forall u, v \in V: f_G(u, v) \leq c(u, v)$;
- антисимметричность: $\forall u, v \in V: f_G(u, v) = -f_G(v, u)$;
- сохранение потока: $\forall u \in V \setminus \{s, t\}: \sum_{v \in V} f_G(u, v) = 0$.

Через функцию $|f_G|$ обозначается величина потока в графе G . Она показывает количество информации, передаваемое из источника в сток в единицу времени. Функция $|f_G|$ вычисляется следующим образом:

$$|f_G| = \sum_{v \in V} f_G(s, v). \quad (4)$$

Максимальную величину потока в графе G среди всех возможных значений f_G обозначают $|F_G|$.

Возможным планом аренды новых каналов является множество ребер E^* такое, что $E^* \subseteq E'$. В случае одобрения реализации плана проектирования исходная сеть преобразуется в сеть, что описывается графом $G^* = (V, E \cup E^*)$.

2.3 Задача оптимизации сети передачи данных

В качестве критериев оптимизации рассматривается стоимость аренды каналов

$$Q_1(E^*) = \sum_{(u,v) \in E^*} p'(u, v) \quad (5)$$

и величина максимальной пропускной способности сети, взятая со знаком минус.

$$Q_2(E^*) = \sum_{v \in V} f(s, v). \quad (6)$$

Задача оптимизации формулируется следующим образом:

По заданному ациклическому ориентированному графу $G = (V, E)$, множеству ребер E' ($E \cap E' = \emptyset$), матрицам пропускных способностей $c(u, v)$ и $c'(u, v)$ и матрице стоимости аренды ребер $p'(u, v)$ найти множество Парето-оптимальных решений задачи бикритериальной оптимизации

$$\min_{E^*}(Q_1(E^*)), \min_{E^*}(Q_2(E^*)). \quad (7)$$

Выбор данных критериев оптимизации имеет ряд причин. В реализации любого проекта стоимость является важнейшим фактором, а величина арендной платы, с одной стороны, выступает как ограничение в виде бюджета, выделенного на модернизацию, а с другой – как критерий оптимизации.

Оценить предельное количество пользователей сети помогает величина максимальной пропускной способности модифицированной СПД. Эта информация полезна при проектировании новых отделений банка, точек самообслуживания клиентов и т.п. Таким образом, банк стремится к качественному и быстрому обслуживанию клиентов.

Найденное множество Парето-оптимальных решений позволит проектировщикам отдела систем связи и телекоммуникаций выбирать между более высокими затратами на модернизацию, получая максимальную пропускную способность сети, и значительно меньшими, получая соответственно меньшую пропускную способность.

Далее рассматриваются существующие методы решения задачи оптимизации СПД, описанные в работах [2, 5, 6, 10, 16, 18, 19, 31, 37].

2.4 Определение множества Парето-оптимальных решений задачи оптимизации сети передачи данных путем полного перебора

Данный метод представляет собой полный перебор всех возможных вариантов структуры СПД, где для каждой сети рассчитывается максимальная пропускная способность и стоимость аренды новых каналов.

Для поставленной задачи (5) доказано, что множество оптимальных по Парето решений может состоять из всех элементов множества решений D , мощность которого экспоненциально зависит от размерности задачи (количества ребер E'); доказана NP-трудность задачи (5) нахождения всего множества оптимальных по Парето решений [29]. Таким образом мощность множества допустимых решений D равна $2^{|E'|}$.

Последний улучшенный алгоритм поиска максимального потока – алгоритм Орлина [40]. Используя данный алгоритм для нахождения максимальной пропускной способности сети, можно добиться времени выполнения равной $O(|V| \cdot |E|)$.

Тогда, время работы алгоритма составит порядка $O(2^{|E'|} \cdot |V| \cdot |E|)$.

В статье [17] приведены результаты тестирования алгоритма полного перебора. Данный алгоритм не имеет практической пользы в решении рассматриваемой задачи, так как время, за которое он находит множество Парето, неприемлемо велико. Однако его реализация играет важную роль в оценке работы других, более сложных методов решения.

2.5 Определение множества Парето-оптимальных решений задачи оптимизации сети передачи данных методом ветвей и границ

Под методом ветвей и границ понимается алгоритм решения задачи, имеющий древовидную структуру поиска оптимального решения и использующий результаты решения оценочных задач [31]. Древовидная

структура, описывающая разбиение всего множества решений на непересекающиеся подмножества D_1, D_2, \dots, D_n ($D = D_1 \cup D_2 \cup \dots \cup D_n$, $D_i \cap D_j = \emptyset, i \neq j$), называется деревом ветвления. Далее приведено описание метода ветвей и границ для решения задачи МКО (5).

1) Все ребра из множества E' нумеруются произвольным образом различными числами от 1 до $|E'|$. Через e_i обозначается ребро с номером i из множества E' .

2) Каждому подмножеству области допустимых решений D соответствует нижняя (L_p) и верхняя (H_p) оценки значения пары $(Q_1(E^*), Q_2(E^*))$ на этом подмножестве. Корню дерева ветвлений (уровню $k = 0$) соответствует вся область допустимых решений D , а каждой вершине $k > 0$ уровня – подмножество решений, получающееся при фиксировании первых k ребер (из множества E') сети. Под «фиксированием» понимается тот факт, что для выбранных ребер уже принято решение: будут они арендованы или нет. Через S обозначено множество фиксированных ребер, которые будут арендованы.

3) Каждый следующий уровень вершин получается путем дихотомического ветвления из каждой вершины предыдущего уровня. При этом одна из вершин нового уровня соответствует подмножеству, полученному после фиксирования очередного ребра, но без добавления его в S : данное ребро не будет арендовано. Вторая – подмножеству, полученному после фиксирования очередного ребра с добавлением его в S : данное ребро будет арендовано.

4) Нижняя оценка L_p решения E^* после фиксирования первых k ребер вычисляется следующим образом:

$$Q_1(E^*) = \sum_{e \in S} p'(e) + \sum_{i=k-1}^{|E'|} p'(e_i), \quad (8)$$

$$Q_2(E^*) = |F_{G(V, E \cup S)}|, \quad (9)$$

то есть первый критерий нижней оценки L_p равен сумме стоимостей аренды фиксированных ребер, которые будут арендованы, и ребер, которые еще не рассматривались, а второй – величине потока для графа, полученного путем добавления в исходный граф множества арендованных фиксированных ребер.

5) Верхняя оценка H_p решения E^* после фиксирования первых k ребер вычисляется с помощью формул:

$$Q_1(E^*) = \sum_{e \in S} p'(e), \quad (10)$$

$$Q_2(E^*) = |F_{G(V, E \cup S \cup \{e_i | i > k\})}|, \quad (11)$$

то есть первый критерий верхней оценки H_p равен стоимости фиксированных ребер, которые будут арендованы, а второй – величине потока для графа, полученного путем добавления в исходный множества арендованных фиксированных ребер и всех ребер, которые еще не рассматривались.

6) После расчета верхней H_p и нижней L_p оценок решения p , где оценка $(Q_1(r), Q_2(r))$ соответствует решению $r \in R$ (R – множество рекордов), происходит проверка на выполнение следующих условий:

- при $L_p > (Q_1(r), Q_2(r))$ решение $r \in R$ удаляется из множества рекордов, поскольку существует решение его доминирующее;

- при $(Q_1(r), Q_2(r)) > H_p$, ветвление из текущей вершины прекращается и происходит переход к рассмотрению других подмножеств решений;

- если $\nexists r \in R: (Q_1(r), Q_2(r)) > H_p$, то в случае рассмотрения вершины дерева, являющейся листом, решение p добавляется в множество рекордов R ($R = R \cup \{p\}$). В противном случае из текущей вершины продолжается ветвление. То есть, если ни одна нижняя оценка рекорда из R не доминирует верхнюю оценку подмножества решений D_p , и еще не все ребра зафиксированы,

ветвление продолжается, иначе решение p является оптимальным по Парето на всех рассмотренных на данный момент подмножествах и добавляется в R .

Символ \succ обозначает Парето-доминируемость одного решения над другим.

7) Останов алгоритма происходит, когда дальнейшее ветвление осуществить невозможно – после того как были рассмотрены все вершины дерева ветвлений, для которых было рационально выполнить операцию ветвления.

В результате множество R содержит все решения задачи (5), оптимальные по Парето. Стоит отметить, что во множество рекордов попадают исключительно те решения, в которых зафиксированы все ребра.

Одним из главных достоинств предложенного метода является простота реализации. Так же алгоритм не требует явного хранения в памяти подмножеств решений, подвергаемых процедуре ветвления.

В некоторых случаях метод ветвей и границ в процессе работы не отбросит ни одного решения, тогда время выполнения будет то же, что и у метода полного перебора. В работе [17] приведен пример такой сети, в которой множество оптимальных по Парето решений имеет мощность $2^{|E'|}$. В процессе решения будет построено дерево ветвлений, состоящее из $|E'|$ уровней (корень дерева имеет уровень 0). При этом уровень глубины i будет иметь 2^i вершин.

Несмотря на это, среднее время выполнения метода гораздо меньше, чем у алгоритма полного перебора, что доказывается результатами вычислительных экспериментов, приведенными в работе [19]. На рисунке 2 показан график зависимости времени работы алгоритмов полного перебора и метода ветвей и границ от размерности задачи.

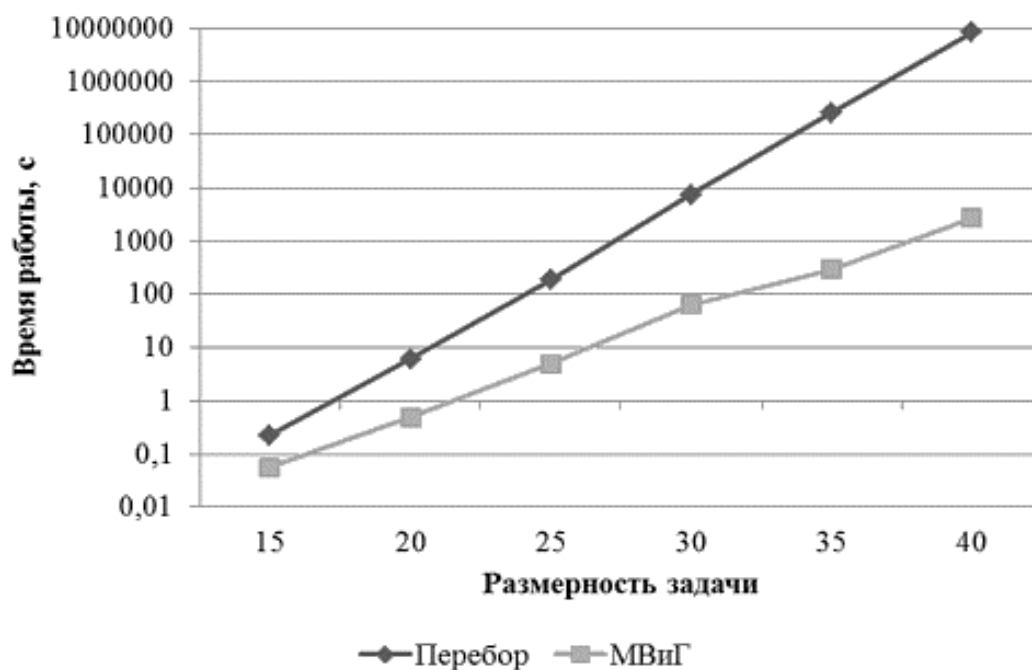


Рисунок 2 – Среднее время работы алгоритма полного перебора и метода ветвей и границ на задачах различной размерности

Количество памяти, необходимое для работы алгоритма, экспоненциально зависит от размера данных, подаваемых на вход. Найденные недоминируемые решения могут достигать величины $2^{|E'|}$. Таким образом, для стабильной работы потребуется количество памяти порядка $O(2^{|E'|} \cdot |E'|)$.

Однако чаще всего множество оптимальных по Парето решений имеет порядок нескольких десятков. Результаты вычислительных экспериментов показали, что для сетей, в которых предлагается к аренде 35 каналов, Парето-совокупность содержит в среднем 41 решение (величина получена при анализе 100 различных примеров сетей). Поэтому на практике требуемое количество памяти невелико.

2.5.1 Пример решения задачи методом ветвей и границ

На рисунке 3 задан граф G и множество ребер, предложенных к аренде E' . Сплошная линия на графе G – это арендованные каналы с указанием пропускной

способности каждого из них, пунктирная линия – каналы из множества E' с указанием пропускной способности и стоимости аренды в месяц для каждого из них. Необходимо найти все множество Парето-оптимальных решений рассматриваемой бикритериальной задачи оптимизации.

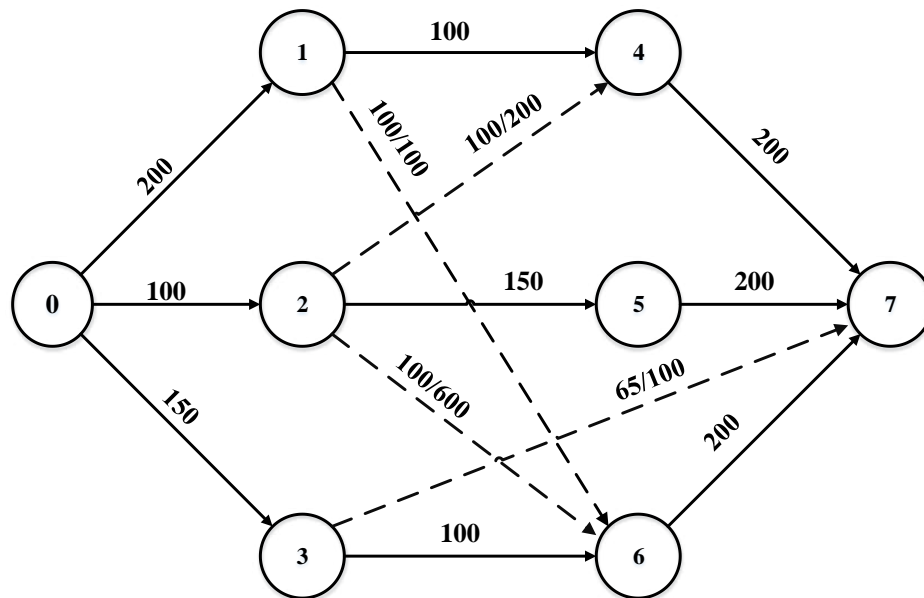


Рисунок 3 – Входные данные для примера

Пронумеруем ребра из E' по порядку, начиная с 1: (1,6), (2,4), (2,6), (3,7), где (u, v) обозначает ребро из вершины u в вершину v . Построим дерево ветвлений (рисунок 4), записывая подмножества решений в таблицу 1.

Второй и третий столбцы таблицы 1 содержат нижнюю и верхнюю оценки подмножества решений D_i , получаемые фиксированием всех ребер из E' при переходе от корня дерева ветвлений до вершины с номером i . Четвертый столбец содержит множество рекордов, найденное на момент входа в вершину i дерева ветвлений.

На ребрах дерева ветвлений указаны числа 0 и 1, показывающие будет ли арендовано рассматриваемое зафиксированное ребро из множества E' или нет соответственно.

Поиск Парето-оптимальных решений начинается в корне дерева ветвлений (вершина 0). Делается предположение, что первое ребро из множества E' не

будет арендовано. Таким образом фиксируется ребро с номером 1 из E' и выполняется переход в дереве ветвлений из вершины 0 в вершину 1 по ребру с 0. Вычисляется нижняя и верхняя оценка для подмножества решений D_1 . Нижняя оценка представляет собой значение стоимости аренды и пропускной способности сети, которые вычисляются следующим образом. Стоимость аренды получается из расчета, что все незафиксированные ребра (с номерами 2, 3, 4) достраиваются, а пропускная способность из расчета, что они не изменяют ее, то есть можно считать, что их пропускные способности равны 0. В результате получаем стоимость аренды, равную $p'(2,4) + p'(2,6) + p'(3,7) = 900$, и пропускную способность, равную 280. При вычислении стоимости аренды для верхней оценки считаем, что ни одно из незафиксированных ребер не арендовано, а при вычислении пропускной способности считаем, что они арендованы.

Таким образом, стоимость аренды равна 0, а пропускная способность равна 345. Далее проверяется, доминируется ли полученная верхняя оценка нижней оценкой элемента из совокупности рекордов R . Так как R пусто, то условие не выполняется, и так как не все ребра из множества E' зафиксированы, то выполняется ветвление из вершины 1 дерева ветвлений.

Аналогично выполняется дальнейшее ветвление, пока все ребра из E' не будут зафиксированы или полученная верхняя оценка не будет доминироваться одним из элементов совокупности рекордов R .

В случае, когда не все ребра из E' зафиксированы, а верхняя оценка доминируется элементом из R , дальнейшее ветвление не имеет смысла, происходит переход в предыдущую вершину ветвления. Такие случаи наблюдаются в вершинах 6, 7, 13 и 14 дерева ветвления.

Если верхняя оценка доминирует элемент из R , то этот элемент удаляется из рекордов, как, например, в вершине 11 (рисунок 4).

В результате решения данного примера найдено множество Парето-оптимальных решений $\{\{\emptyset\}, \{1\}, \{1,4\}\}$, которым соответствуют оценки $(0, -280)$, $(100, -380)$ и $(200, -445)$.

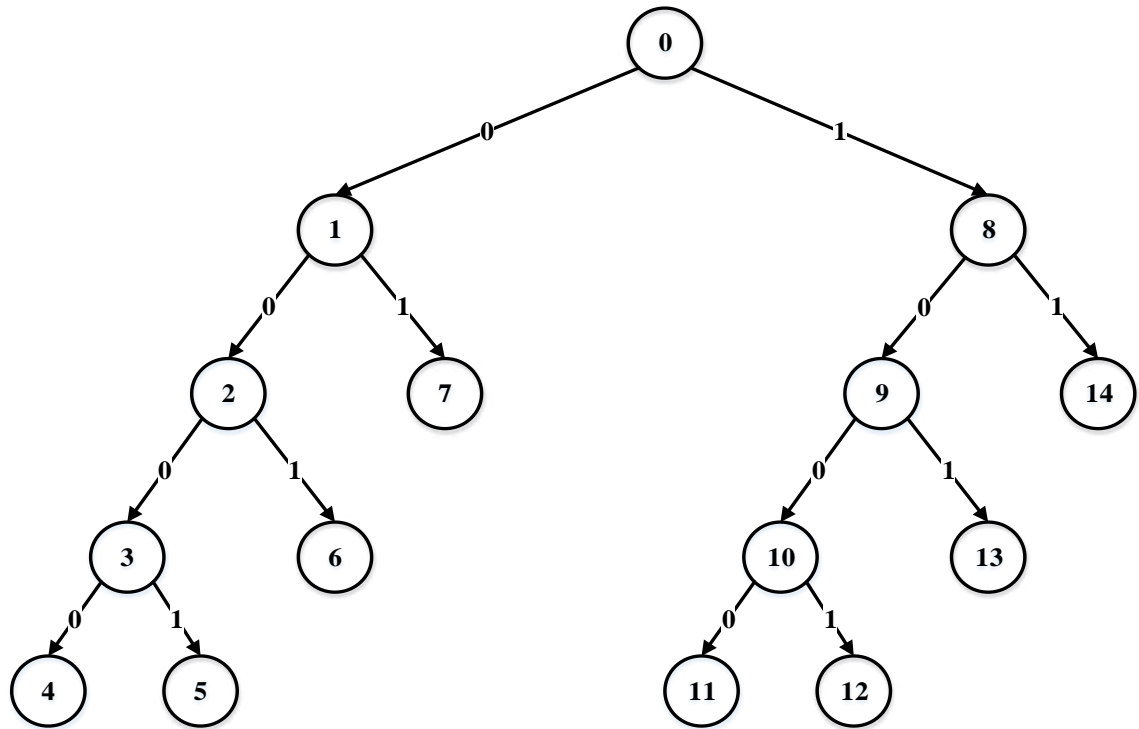


Рисунок 4 – Дерево решений метода ветвей и границ

Таблица 1 – Оценки для подмножеств дерева решений (рисунок 4)

№ вершины	Нижняя оценка D_i	Верхняя оценка D_i	Множество рекордов R
1	(900, -280)	(0, -345)	\emptyset
2	(700, -280)	(0, -345)	\emptyset
3	(100, -280)	(0, -345)	\emptyset
4	(0, -280)	(0, -280)	\emptyset
5	(100, -375)	(100, -345)	(0, -280)
6	(700, -280)	(600, -345)	(0, -280), (100, -345)
7	(900, -280)	(200, -345)	(0, -280), (100, -345)
8	(1000, -380)	(100, -445)	(0, -280), (100, -345)
9	(800, -380)	(100, -445)	(0, -280), (100, -345)
10	(200, -380)	(100, -445)	(0, -280), (100, -345)
11	(100, -380)	(100, -380)	(0, -280), (100, -345)
12	(200, -445)	(200, -445)	(0, -280), (100, -380)
13	(800, -380)	(700, -445)	(0, -280), (100, -380), (200, -445)
14	(1000, -380)	(300, -445)	(0, -280), (100, -380), (200, -445)

2.6 Выводы по главе 2

Построена математическая модель СПД в виде взвешенного ориентированного ациклического графа, которая учитывает особенности сети Сибирского банка.

В рамках построенной модели представлена бикритериальная задача оптимизации СПД, проведен анализ ее вычислительной сложности.

3 Поиск субоптимального множества решений задачи оптимизации сети передачи данных

Число Парето-оптимальных решений может экспоненциально зависеть от размерности задачи (5), а значит множество Парето может оказаться слишком большим, что затруднит или сделает невозможным анализ результатов для ЛПР. Тогда поиск всего множества Парето становится нецелесообразным.

В этом случае вводится понятие субоптимального множества.

3.1 Субоптимальное множество

Задача (5) принадлежит к классу NP-трудных, поэтому решить её с помощью полиномиальных алгоритмов не представляется возможным, учитывая гипотезу о $P \neq NP$. В реальных условиях количество предложенных к аренде каналов может быть около сотни и более, тогда время работы такого алгоритма станет слишком велико.

В таком случае активно применяются эвристические методы, которые возможно найти не точное, но приближенное к оптимальному решение, применив эвристические методы. При этом вычислительная сложность данных алгоритмов позволяет находить ответ к задаче за приемлемое время.

Доминируемые решения, найденные эвристическим алгоритмом, не представляют интереса и поэтому могут быть исключены из рассмотрения. В связи с этим дадим следующее определение.

Множество решений P' ($\forall p \in P': p \in D$) называется субоптимальным при условии, что каждый элемент из P не доминирует и не равен ни одному другому элементу из множества P .

Таким образом, в условиях большой размерности задачи (5) целесообразно найти субоптимальное множество решений, что позволит упростить анализ результата для ЛПР.

3.2 Оценка отклонения субоптимального множества решений от множества Парето

Под эффективностью алгоритма зачастую подразумевается качество результата и требуемые вычислительные ресурсы.

В отношении ресурсов используются фиксированное число операций вычисления или константное время выполнения в общем. В таком случае, нет разницы между однокритериальной и многокритериальной оптимизацией. Серьезная разница наблюдается при оценке качества получаемого результата. В однокритериальной оптимизации качество решения может определяться с помощью целевой функции. Однако в случае МКО не понятно, что значит качество – близость к фронту Парето, охват как можно большего количества разнородных решений или какие-то другие признаки.

В работах [8, 20] предложены варианты оценки качества найденного субоптимального множества.

3.2.1 Метод подсчета решений

Метод оценивает качество полученного субоптимального множества решений по отношению между количеством «плохих» или ненайденных оптимальных решений и количеством элементов множества Парето-оптимальных решений задачи.

В данном контексте «плохое» решение – это такое решение x , которое входит в субоптимальное множество P' , но не принадлежит множеству Парето-оптимальных решений P поставленной задачи. Количество таких решений находится по формуле:

$$Q_b = |\{x: x \in P', x \notin P\}|, \quad (12)$$

Тогда поиск отклонения осуществляется с помощью формулы:

$$\Delta_1(P, P') = \frac{\max(Q_b, Q_p)}{|P|}, \quad (13)$$

где Q_p – количество Парето-оптимальных решений, не найденных алгоритмом, P – количество Парето-оптимальных решений.

Соответственно, чем больше значение Δ_1 , тем хуже результат.

Основным недостатком метода является то, что в нем не учитывается точность приближения субоптимального множества к множеству Парето-оптимальных решений. Поэтому рекомендуется дополнительно применять описанный ниже метод.

3.2.2 Метод оценки точности приближения субоптимального множества решений к Парето-оптимальному

Точность приближения вычисляется следующим образом: для каждой оценки решения из субоптимального множества находится отношение минимального расстояния до множества Парето к расстоянию между оценкой ближайшего Парето-оптимального решения и оценкой начального состояния сети. Под начальным состоянием сети подразумевается текущее состояние используемой в настоящий момент сети. В качестве оценки точности приближения берется среднее значение найденных отношений. Для вычисления расстояния между оценками используется метрика Чебышева:

$$\rho(x; x') = \max(|Q_1(x) - Q_1(x')|, |Q_2(x) - Q_2(x')|). \quad (14)$$

Формула оценки точности приближения представлена ниже.

$$\Delta_2(P, P') = \underset{(Q_1(x'), Q_2(x')) \in P'}{avg} \left(\frac{\min_{(Q_1(x), Q_2(x)) \in P} \rho(x; x')}{\rho(x''; x_0)} \right), \quad (15)$$

где $\rho(x_1; x_2)$ – расстояние между двумя оценками $(Q_1(x_1), Q_2(x_1))$ и $(Q_1(x_2), Q_2(x_2))$ решений x_1 и x_2 соответственно, вычисляемое с помощью метрики Чебышева;

x_0 – решение, соответствующее начальному состоянию сети;

x – Парето-оптимальное решение;

x' – субоптимальное решение;

x'' – решение из множества Парето, ближайшее к субоптимальному x' .

Этот метод оценки не учитывает наличие «плохих» решений. Чтобы избежать случаев, когда оценка субоптимального множества показывает большой процент таких решений, рекомендуется совместное применение оценок (13) и (15), поскольку это дает более полное представление о качестве и допустимости субоптимальных решений.

3.3 Поиск субоптимального множества решений с применением имитации отжига

Метод имитации отжига представляет собой алгоритмический аналог физического процесса образования в веществе кристаллической структуры, характеризующейся минимальной энергией, при управляемом охлаждении. Другими словами, данный метод использует упорядоченный случайный поиск для определения новых состояний системы.

Несмотря на широкую область применения: от решения оптимизационных задач до обучения нейронных сетей, скорость сходимости метода остается мало изученной. Однако в работе Л. Инберга [38] показана эффективность метода для большого класса задач. В сравнении же с генетическими алгоритмами [39] данный метод не только не хуже, но и во многих случаях лучше.

Алгоритм имитации отжига для нахождения экстремума функции состоит из следующих основных шагов:

- 1) создание начального вектора решений S ;
- 2) изменение текущего решения случайным образом, вычисление энергии;
- 3) проверка критерия допуска, замена текущего решения новым по необходимости.
- 4) снижение температуры;

5) проверка критерия останова, по результату завершение или переход к шагу 2.

Далее подробно рассмотрены основные моменты алгоритма имитации отжига.

3.3.1 Начальный вектор решений. Получение нового решения

Начальное решение представляет собой вектор решений S , каждый элемент которого будет независимо от других подвержен процедуре отжига. Это позволит с большей вероятностью найти Парето-оптимальные решения, значительно отличающиеся друг от друга.

Положим размер вектора решений S равным $|E'|^2$, а все элементы множества E' пронумерованы по порядку. Решение S_i представляет собой упорядоченный вектор размера $|E'|$, состоящий из нулей и единиц, которые кодируют состояние системы. То есть, состояние равно 1, если соответствующий ребру e_j ($e_j \in E'$) канал арендован, и 0 в противном случае.

Генерация начального решения из вектора S происходит путем выбора произвольного целого числа k от 0 до $|E'|$, и заполнения k произвольных элементов вектора 1, а остальных 0.

Поскольку метод имитации отжига подразумевает переход из текущего состояния в одно из «соседних» - схожего по структуре решения, достаточно случайным образом выбрать $\left\lceil \frac{|E'|}{8} \right\rceil$ ($\lceil k \rceil$ – операция округления числа k в большую сторону) компонентов вектора и произвести их инверсию (0 заменить 1 и наоборот).

3.3.2 Снижение температуры. Вычисление энергии.

Для начального момента времени выбрана температура $T_{max}=100$. Далее после каждой десятой итерации ($t=10$) происходит снижение температуры согласно закону:

$$T_{i+1} = cT_i, \quad (16)$$

где коэффициент $c = 0.98$ [23].

Для расчета энергии перехода из состояния x с оценкой $(Q_1(x), Q_2(x))$, в состояние x' с оценкой $(Q_1(x'), Q_2(x'))$, чаще используют формулу:

$$e^{-\left(\frac{\max(|Q_1(x)-Q_1(x')|, |Q_2(x)-Q_2(x')|)}{T}\right)}, \quad (17)$$

3.3.3 Критерии допуска и критерий останова

Замена текущего решения на новое происходит, если выполняется хотя бы одно из условий (критериев допуска):

- 1) Значение энергии, рассчитанное по формуле (16) не превосходит некоторой случайной величины, выбранной равновероятно на отрезке $[0,1]$.
- 2) $Q_1(x) - Q_1(x') < 0$, то есть новое решение x' имеет меньшую стоимость аренды соответствующей СПД.
- 3) $Q_2(x) - Q_2(x') > 0$, то есть новое решение x' имеет большую пропускную способность соответствующей СПД.

Кроме того, при выполнении хотя бы одного из описанных критериев, решение x' становится кандидатом на добавление во множество недоминируемых решений NS . Иначе, рассматриваемое решение заменяется на произвольное из S .

Алгоритм прекращает работу после очередного снижения температуры до определенного минимального значения $T_{min} = 0.5$.

3.3.4 Дополнительные эвристики

Поскольку на практике решения с небольшими, с точки зрения современных вычислительных способностей, различиями в общей пропускной способности сети и, одновременно, стоимости аренды не представляют интереса, предлагается ввести настраиваемую ЛПР погрешность при проверке на доминируемость.

Например, множество недоминируемых оценок решений содержит элемент (1028, 1000), который проверяется на доминируемость оценкой (1029, 1030). ЛПР считает несущественным изменение пропускной способности сети в пределах 500 Мб и стоимости аренды в 300 у.е.. Тогда решение (1028, 1000) заменяется на доминирующее его (1029, 1030). Таким образом можно при необходимости сократить количество недоминируемых решений, исключив однообразные по мнению ЛПР решения, тем самым облегчив дальнейший анализ.

3.3.5 Оценка вычислительной сложности

Используя новейший алгоритм Орлина [40] для нахождения максимального потока время поиска максимальной пропускной способности сети и стоимости аренды каналов будет порядка $O(|V| \cdot |E|)$. Количество итераций при снижении температуры от T_{max} до T_{min} по закону (16) равно $t \cdot \log_c \left(\frac{T_{min}}{T_{max}} \right)$. Поиск нового решения происходит за время $O(|E'|)$.

Итого, на вычисление вектора решений S на каждой итерации потребуется порядка $O(|E'| \cdot (|V| \cdot |E| + |E'|))$ времени. Всего количество итераций равно величине $|S| \cdot t \cdot \log_c \left(\frac{T_{min}}{T_{max}} \right)$.

3.4 Комбинированный метод

Использование эвристик на определенных этапах работы метода ветвей и границ позволяет сократить время работы алгоритма, при этом не теряя в качестве результата.

В работе [21] предложено использовать эвристику на этапе инициализации множества рекордов R , что позволяет на начальных итерациях отсеять заведомо неоптимальные подмножества решений.

3.5 Вычислительный эксперимент

Эксперимент проводился на компьютере с четырехъядерным процессором Intel® Core™ i7-4710MQ, частота которого равна 2.5 ГГц, и 12 Гб оперативной памятью DDR4.

В процессе эксперимента были сгенерированы случайным образом 100 различных СПД с $|E'| = 15, 20, 25, 30, 35, 40$, для которых было рассчитано среднее время работы реализованных алгоритмов и произведена оценка отклонения субоптимального множества, найденного алгоритмом имитации отжига.

В таблице 2 указаны экспериментальные данные по времени работы реализованных алгоритмов в зависимости от входных данных. На основе этих данных был построен сравнительный график (рисунок 5). На графике видно, что лучший результат при $|E'|$ не превышающем 25 показывает метод ветвей и границ, от 25 до 30 – комбинированный метод, более 30 – алгоритм имитации отжига.

Таблица 2 – Среднее время работы алгоритмов в зависимости от входных данных

$ E' $	$ V $	Метод ветвей и границ			Алгоритм имитации отжига			Комбинированный метод		
		t_{min}, c	t_{max}, c	t_{cp}, c	t_{min}, c	t_{max}, c	t_{cp}, c	t_{min}, c	t_{max}, c	t_{cp}, c
15	13	0,004	0,197	0,050	0,972	9,975	4,115	1,115	6,091	1,537
20	14	0,04	2,29	0,35	4,81	18,76	8,97	1,20	7,42	1,50

25	15	0,10	34,11	6,53	5,59	31,83	18,13	3,22	43,40	7,71
30	17	0,41	220,49	59,62	6,51	94,82	29,70	3,94	293,34	36,02
35	17	0,99	4290,01	279,54	11,08	112,27	46,69	4,06	1286,1	144,21
40	19	410,08	35330	2330,01	28,56	151,71	82,07	20,25	17112	536,61

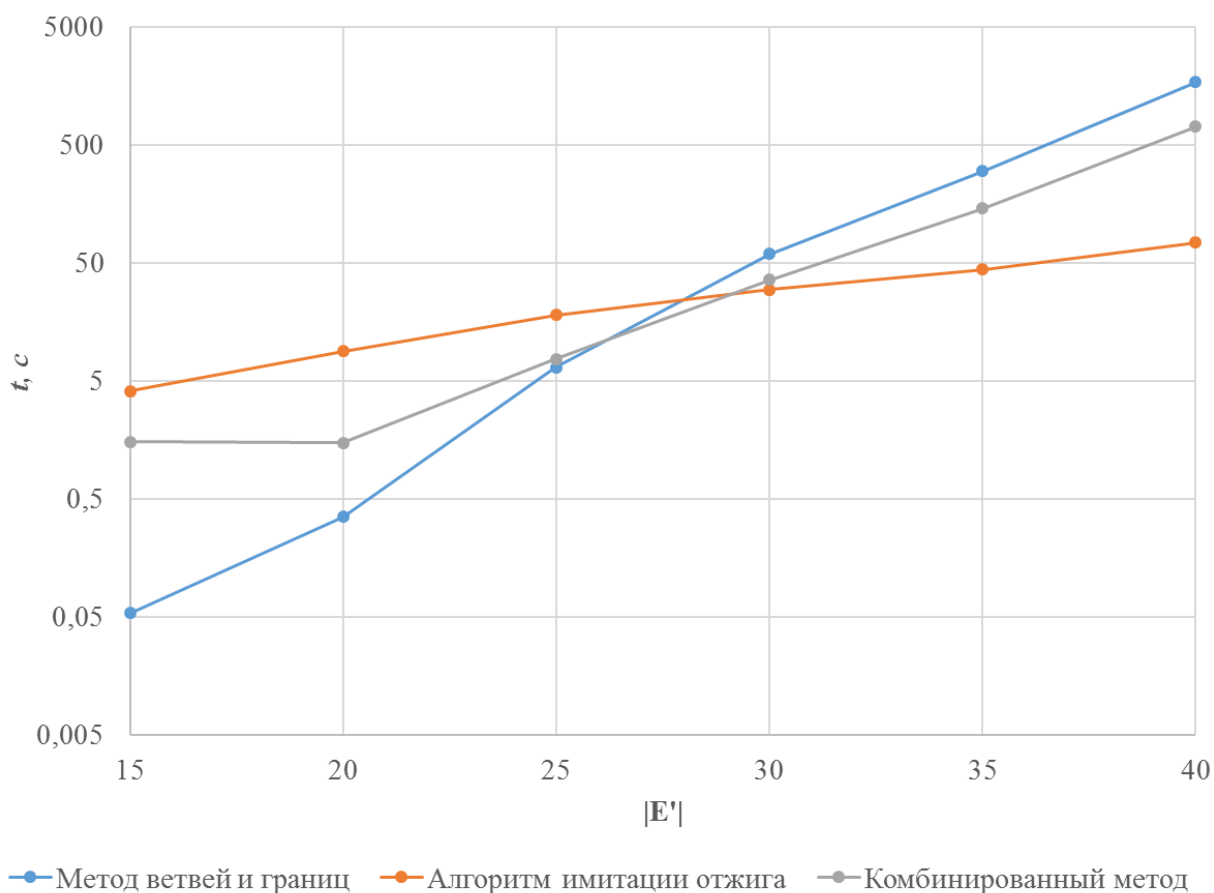


Рисунок 5 – Сравнительный график среднего времени работы алгоритмов

В таблице 3 указаны сведения об отклонении множества субоптимальных решений, найденного с помощью алгоритма имитации отжига, от множества Парето.

Таблица 3 – Оценка отклонения субоптимального множества от множества Парето-оптимальных решений

$ E' $	$ V $	$\Delta_1(P, P')$			$\Delta_2(P, P')$		
		$\Delta_{min}, \%$	$\Delta_{max}, \%$	$\Delta_{cp}, \%$	$\Delta_{min}, \%$	$\Delta_{max}, \%$	$\Delta_{cp}, \%$

15	13	10,01	81,97	48,14	0,00	15,11	6,08
20	14	10,58	90,93	42,76	2,00	44,00	13,70
25	15	21,16	97,57	57,68	3,46	30,60	17,32
30	17	28,15	97,59	64,31	2,99	25,83	13,37

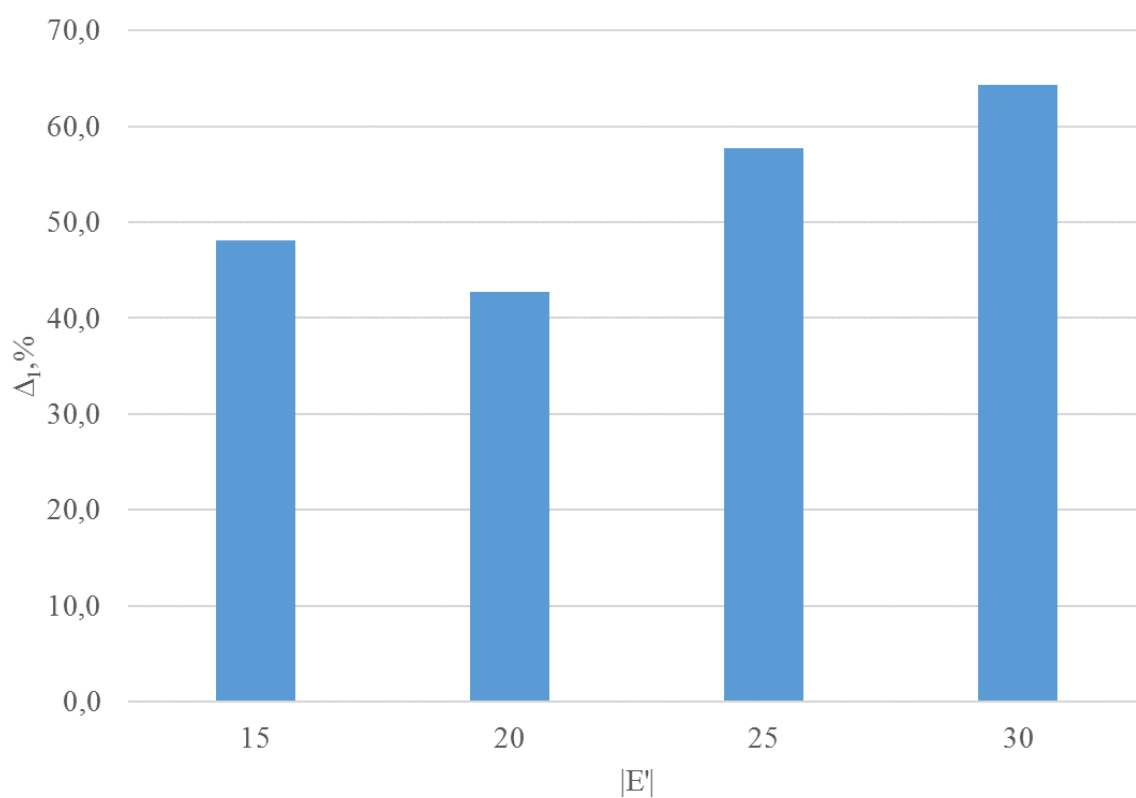


Рисунок 6 – График оценки $\Delta_1(P, P')$

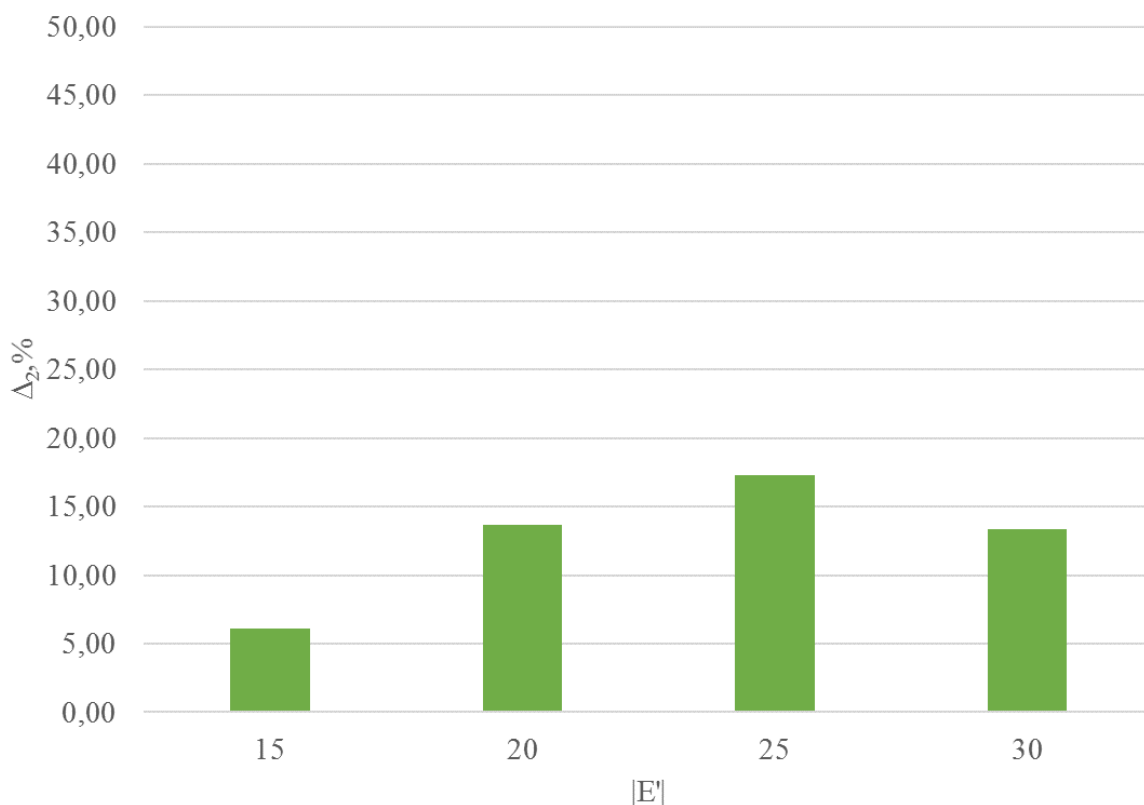


Рисунок 7 – График оценки $\Delta_2(P, P')$

Вычислительный эксперимент показал, что для наискорейшего получения решений теоретических задач использование эвристики целесообразно, когда количество предлагаемых к аренде каналов больше 30. При меньшем количестве метод ветвей и границ успешно справляется с решением.

Однако для решения практических задач рекомендуется отдать предпочтение точным методам, поскольку применение только эвристических методов предполагает риски, которые стоит избегать. Таким образом, при решении практических задач, когда количество предлагаемых к аренде каналов превышает 25 рекомендуется применять комбинированный метод. Его применение позволит получить решение за меньшее, по сравнению с методом ветвей и границ, время, не потеряв при этом в качестве. Кроме того, дополнительная эвристика позволит сократить множество Парето-оптимальных решений, что упростит дальнейший анализ результатов для ЛПР.

4 Программная реализация методов решения задачи оптимизации сети передачи данных

4.1 Краткое описание

Конечным пользователем программы является инженер отдела систем связи и телекоммуникаций Сибирского банка.

Разработанное программное обеспечение позволяет пользователю:

- задавать СПД в графическом и матричном виде;
- сохранять описание сети в файл и загружать из файла;
- формировать и сохранять отчет в формате txt, составленный согласно стандарту организации;
- выбирать метод решения задачи самостоятельно или оставить рекомендуемый для заданной размерности задачи метод.

Интуитивно понятный графический интерфейс программы помогает пользователю быстро и эффективно решать поставленные перед ним задачи с помощью разработанного программного обеспечения.

Программа предусматривает расширенные настройки для опытных пользователей, для более гибкого подхода к решению специфических задач.

4.2 Средства разработки

Программный комплекс реализован в интегрированной среде разработки Microsoft Visual Studio 2012 на языке программирования высокого уровня C# версии 5.0.

Графический пользовательский интерфейс создан с помощью технологии Windows Presentation Foundation (WPF) на языке XAML.

4.3 Рекомендуемые системные требования

Рекомендуемая тактовая частота процессора 2 ГГц или более, от 1 ГБ ОЗУ для Windows 7/8/10. Операционная система Microsoft Windows 7/8/10 32-разрядная или 64-разрядная. Требуется 10 МВ свободной памяти на системном диске;

4.4 Прототип графического пользовательского интерфейса

4.4.1 Вид главного окна программы

На рисунке 8 и 9 представлены скриншоты главного окна программы.

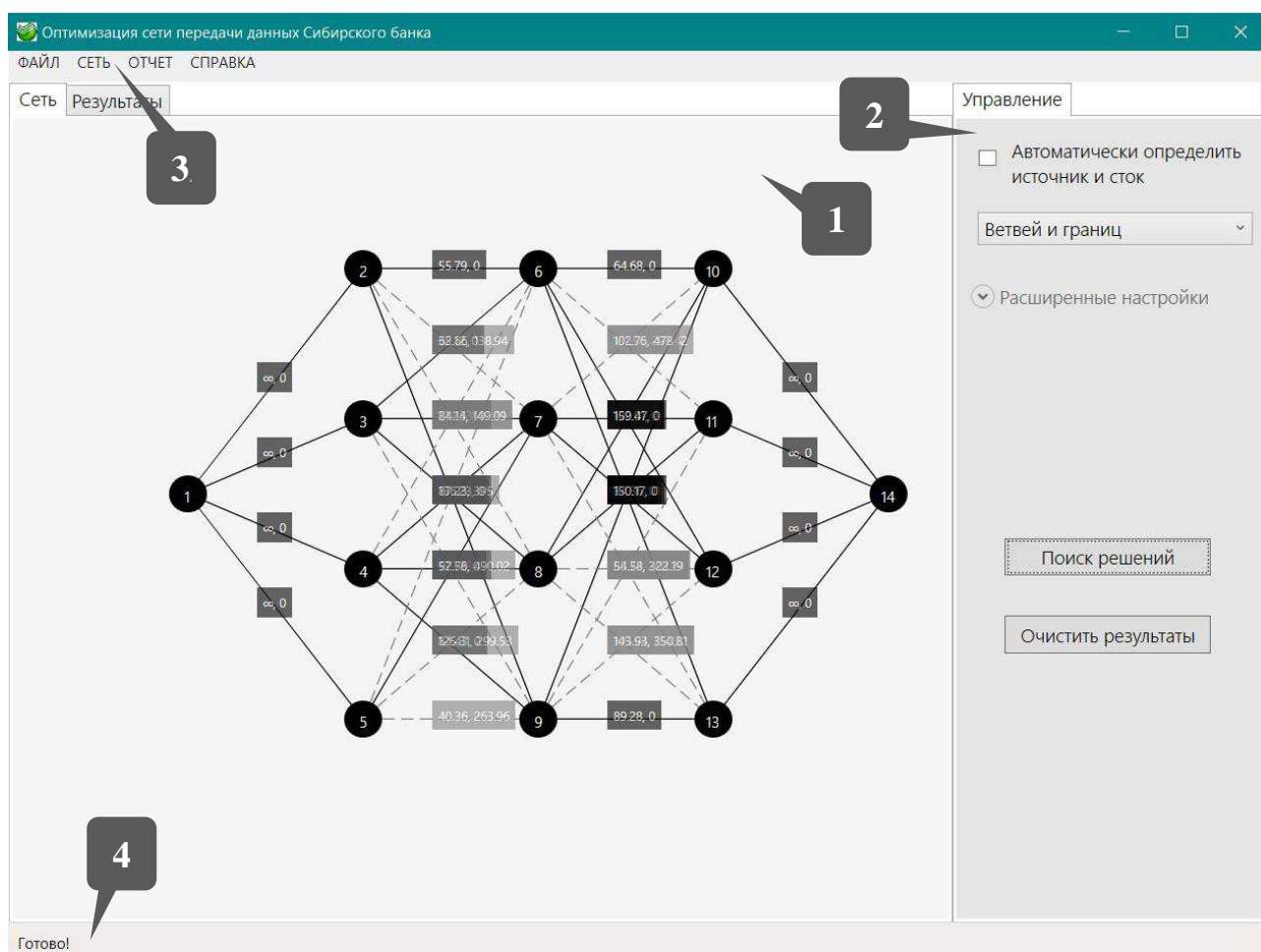


Рисунок 8 – Скриншот главного окна программы

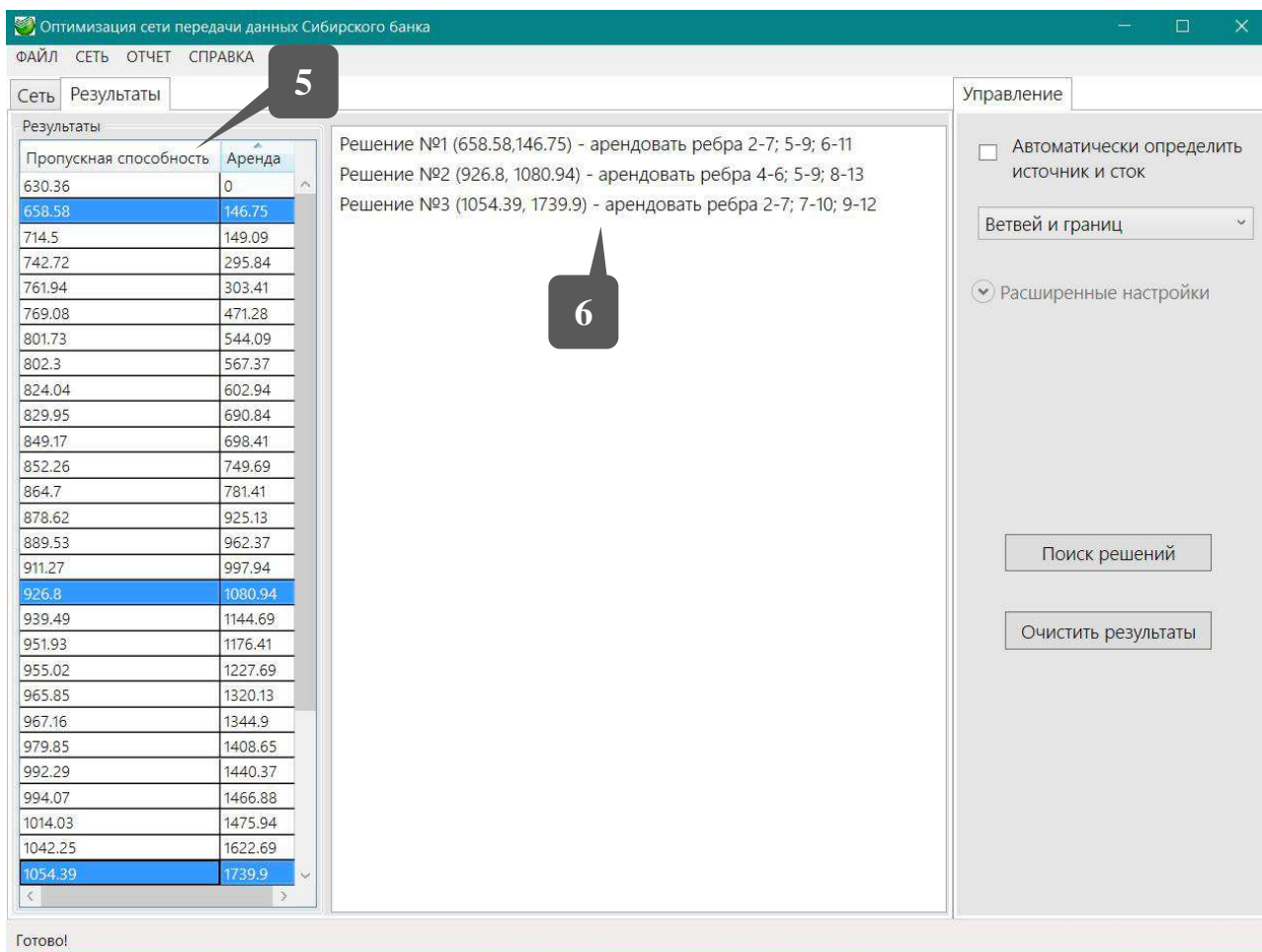


Рисунок 9 – Скриншот главного окна программы с результатами

Ниже приведено описание обозначений на рисунках 8 и 9.

1) Область ввода и отображения СПД. Узловые элементы сети представляют собой окружности с идентификационным номером внутри. Для добавления узла достаточно кликнуть левой кнопкой мыши (ЛКМ) на свободное пространство внутри области. Каналы обозначены линиями, соединяющими соответствующие узлы. Для каждого канала указана пропускная способность и величина арендной платы. Предложенные провайдером каналы (еще не арендованные) обозначены пунктирными линиями серого цвета, уже арендованные – сплошными линиями черного цвета. Для добавления нового канала необходимо выбрать сначала начальную вершину, а затем конечную. Выбор вершины осуществляется кликом ЛКМ на ее изображение. Область

отображения СПД масштабируема, что позволяет просматривать сеть с большим количеством элементов.

2) Область с управляющими элементами. В ней находится: переключатель автоматического режима определения источника и стока в сети; выбор реализованных методов решения задачи оптимизации СПД; расширенные настройки для опытных пользователей; кнопка запуска решения задачи оптимизации заданной СПД и кнопка очистки полученных ранее результатов.

3) Главное меню программного комплекса. Отсюда осуществляется доступ к имеющимся функциональным возможностям, включая информацию о программе и справку.

4) Статусная строка. Показывает подсказки для пользователя, сообщает о статусе выполнения операции.

5) Таблица результатов. Каждая строка соответствует одному решению. В первом столбце указывается пропускная способность сети, во втором – стоимость, на которую увеличится общая арендная плата.

6) Поле просмотра выбранных в таблице результатов решений.

4.4.2 Главное меню

Главное меню состоит из следующих подменю.

Пункты меню Файл представлены на рисунке 10. Пункт меню «Создать» позволяет начать новую задачу оптимизации. «Открыть» позволяет загружать из текстового файла описание СПД с последующей графической интерпретацией. «Сохранить» позволяет сохранить изменения в существующий файл, «Сохранить как...» позволяет сохранить сеть в новый файл. Пункт меню «Выход» закрывает программу.

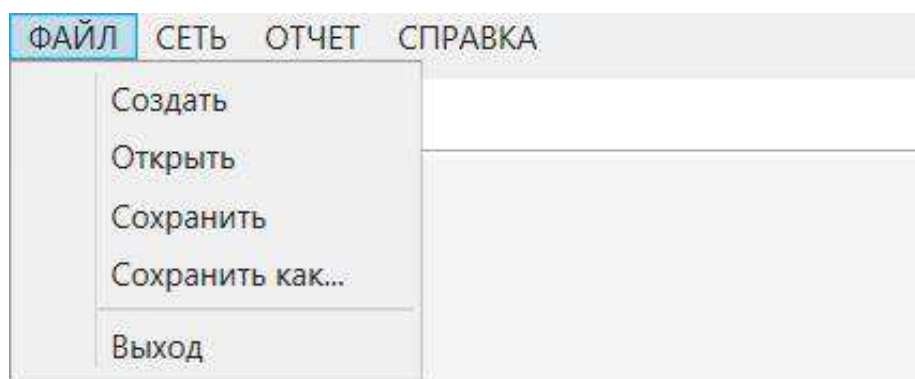


Рисунок 10 – Меню Файл

Меню Сеть представлено на рисунке 11. Пункт «Новый» удаляет все сведения о предыдущей сети и позволяет начать новое построение. «Очистить» удаляет все решения на данном графе. Пункт «Поиск» запускает алгоритм нахождения решений оптимизации СПД. Пункт «Моделирование» позволяет сгенерировать случайным образом топологию сети в соответствии с заданными параметрами.

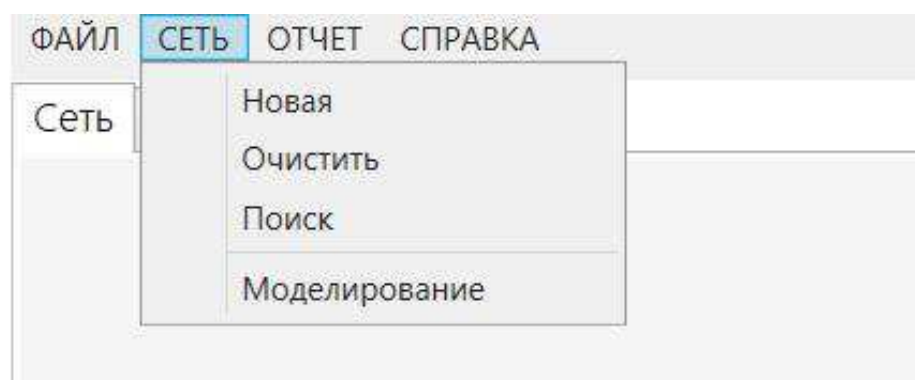


Рисунок 11 – Меню Сеть

Пункты меню Отчет представлены на рисунке 12. Пункт «Сформировать отчет» позволяет оформить результаты в виде документа, оформленного в соответствии со стандартом организации. Пункт «Сохранить» позволяет сохранить сформированный отчет в заданную директорию.

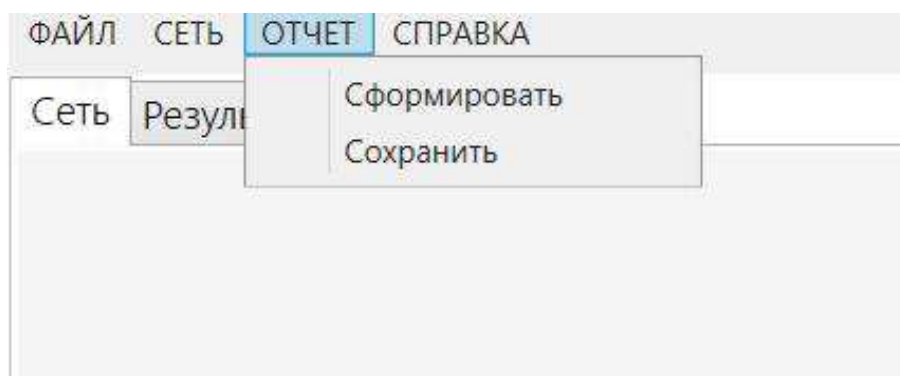


Рисунок 12 – Меню Отчет

Меню Справка представлено на рисунке 13. «О программе» предоставляет основные сведения о версии и разработчике.

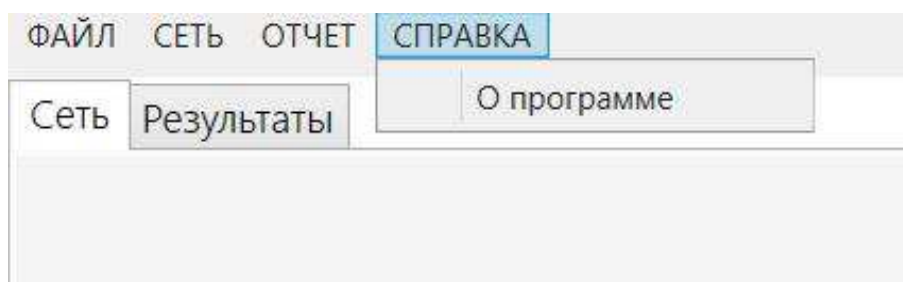
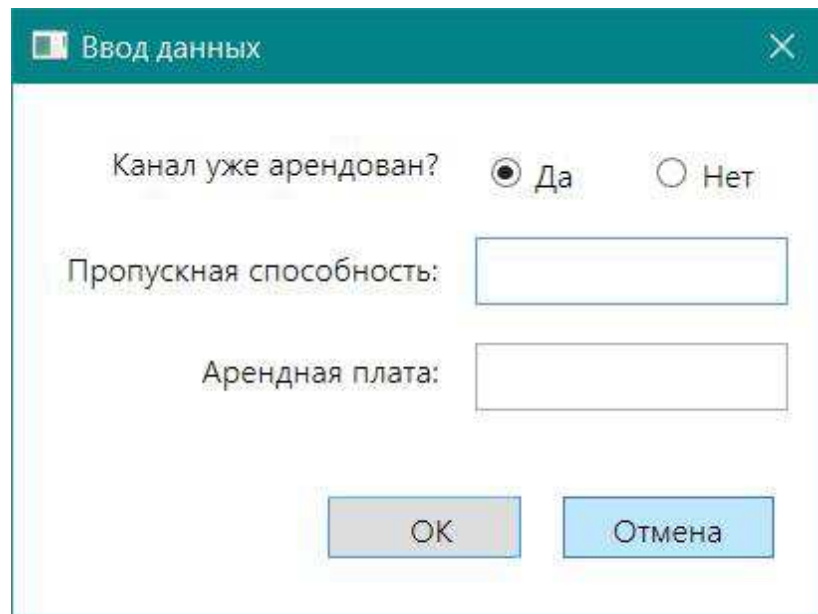


Рисунок 13 – Меню Справка

4.5 Окна ввода данных

На рисунке 14 представлено окно ввода данных о канале связи. В процессе ввода данных заполняются такие поля, как пропускная способность и арендная плата, а также выбирается арендован канал или нет (предлагается провайдером).

На рисунке 15 представлено окно ввода параметров сети для моделирования. В процессе ввода данных заполняются такие поля, как количество коммутаторов, арендованных и доступных к аренде ребер, границы пропускной способности и арендной стоимости каналов, в которых будут сгенерированы случайные значения.



Ввод данных

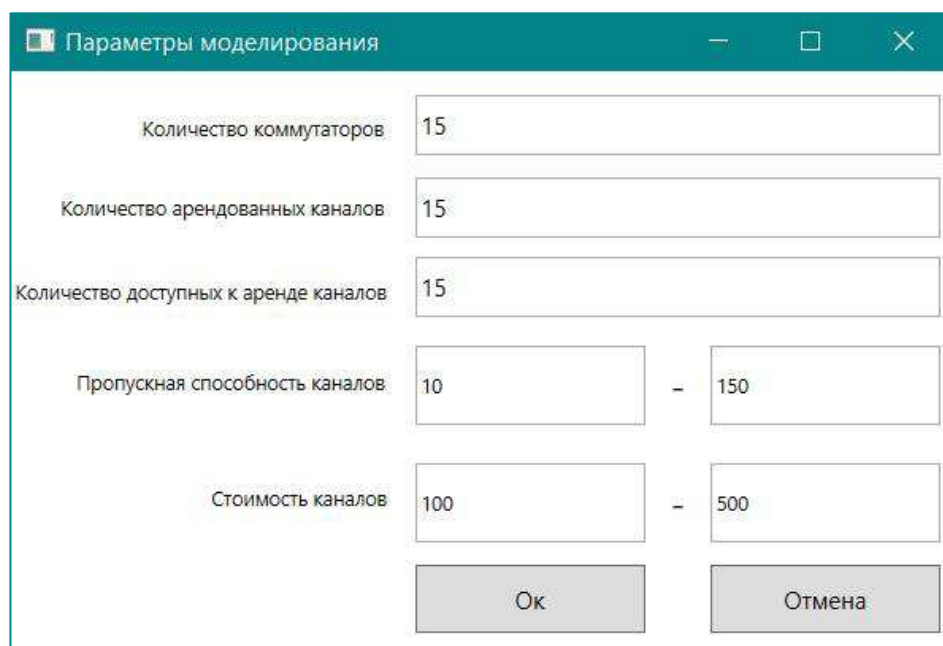
Канал уже арендован? ☒ Да ☐ Нет

Пропускная способность:

Арендная плата:

ОК Отмена

Рисунок 14 – Окно ввода данных о канале связи



Параметры моделирования

Количество коммутаторов

Количество арендованных каналов

Количество доступных к аренде каналов

Пропускная способность каналов –

Стоимость каналов –

Ок Отмена

Рисунок 15 – Окно ввода параметров моделирования

4.6 Типовой сценарий работы с программой

Типовой сценарий работы с программой предполагает следующие действия:

- 1) Задание структуры СПД путем загрузки описания из файла и/или непосредственным добавлением новых вершин и ребер с соответствующими параметрами (пропускной способностью каналов и их арендной стоимостью).
- 2) Запуск алгоритма на выполнение.
- 3) По окончании работы алгоритма (пользователь ориентируется на статус выполнения операции) просмотр найденных решений в таблице.
- 4) Анализ полученных результатов, выбор оптимального для ЛПР плана аренды дополнительных каналов передачи данных.
- 5) Формирование отчета и сохранение его в желаемую директорию.

4.7 Режимы использования программы

4.7.1 Автоматический режим

Автоматический режим представляет собой использование программы со стандартными настройками. В этом случае при решении задачи оптимизации заданной СПД автоматически будут определены вершина источника и стока, согласно анализу размерности задачи будет выбран оптимальный метод решения из списка реализованных. Данный режим рекомендован, как основной.

4.7.2 Пользовательский режим

Пользовательский режим представляет собой использование программы с измененными параметрами работы: задание вершины источника и вершины стока, выбор метода решения задачи, изменение погрешности пропускной способности и/или арендной стоимости, в пределах которой решения являются доминируемыми, задание параметров работы метода имитации отжига и комбинированного метода. Данный режим рекомендуется использовать в том случае, когда результат, полученный в автоматическом режиме по некоторым

причинам не удовлетворяет ЛППР, либо требуется более тонкая настройка работы программы для решения специфической задачи.

4.8 Апробация на реальных данных

4.8.1 Исходные данные

Данные об арендованных каналах связи Сибирского банка на территории Республики Хакасии приведены в таблице Б.1. Со стороны провайдера «Ростелеком» на рассматриваемой территории существует порядка 100 предложений, для апробации программы случайным образом было выбрано 40 из них – это 2^{40} возможных вариантов модернизации. Решение получено комбинированным методом менее, чем за минуту, в то время как у метода ветвей и границ время заняло порядка 10 минут.

4.8.2 Результат работы программы

В результате работы программы получено 9 оптимальных по Парето решений из всего количества возможных вариантов. Кроме того, сформирован отчет с более подробным описанием результатов. Скриншот главного окна с графическим представлением входных данных и окна с результатом работы программы представлены на рисунках 16 и 17.

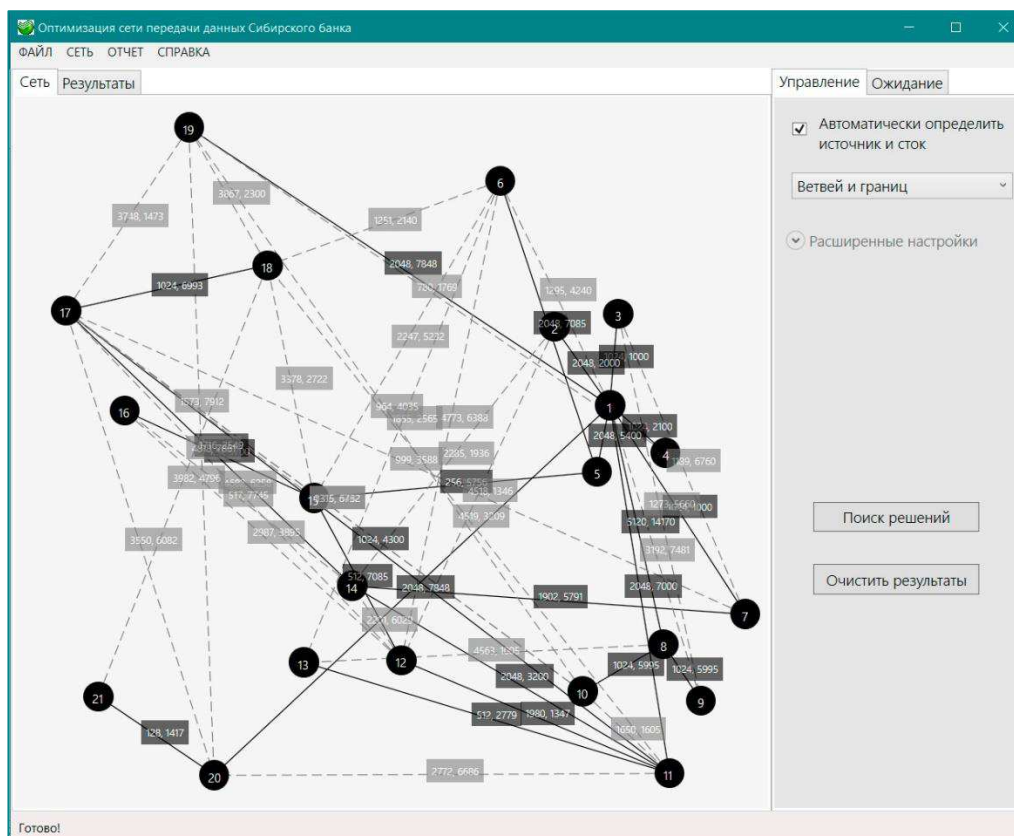


Рисунок 16 – Скриншот главного окна с заданной СПД

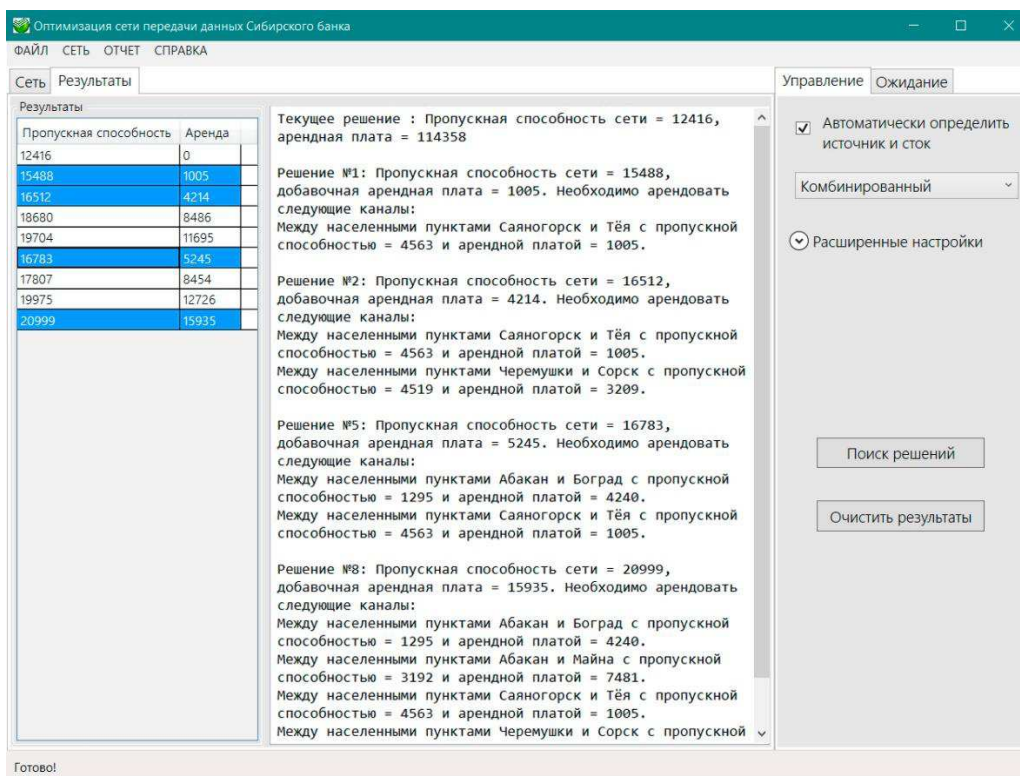


Рисунок 17 – Скриншот главного окна с результатами работы программы

4.8.3 Выводы

В результате работы программы из 2^{40} вариантов было выбрано 9 оптимальных решений, что существенно упрощает работу инженера отдела средств связи и телекоммуникаций при выборе оптимального решения модернизации сети и закладывании ежегодного бюджета.

Сформированный отчет имеет предельно простую структуру представления результатов, что позволяет быстро ориентироваться в тексте и использовать данные для формирования официальных документов.

Апробация работы программы на реальных данных прошла успешно, что подтверждается актом технической экспертизы (приложение А).

ЗАКЛЮЧЕНИЕ

В ходе работы были выполнены следующие задачи:

- построена модель сети передачи данных Сибирского банка;
- поставлена задача бикритериальной оптимизации сети;
- проведен анализ существующих методов решения задач МКО сети передачи данных, выявлены их преимущества и недостатки;
- предложена модификация метода ветвей и границ, позволяющая существенно сократить время работы при решении задачи большой размерности, при этом не потеряв в качестве полученного решения;
- разработана и реализована программа, предназначенная для поиска оптимальных решений оптимизации сети передачи данных Сибирского банка.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Алексеев, О. Г. Комплексное применение методов дискретной оптимизации : науч. изд. / О. Г. Алексеев. – Москва : Наука, 1987. – 248 с.
2. Батищев, Д. И. Генетические алгоритмы решения экстремальных задач : учеб. пособие / Д. И. Батищев. – Воронеж : ВГТУ, 1995. – 62 с.
3. Березовский, Б. А. Многокритериальная оптимизация. Математические аспекты : монография / Б. А. Березовский, Ю. М. Барышников, В. И. Борзенко, Л. М. Кемпнер. – Москва : Наука, 1989. – 128 с.
4. Бертсекас, Д. Сети передачи данных : науч. изд. / Д. Бертсекас, Р. Галлагер. – Москва : Мир, 1989. – 544 с.
5. Бесслер, Р. Проектирование сетей связи : справочник / Р. Бесслер, А. Дойч. – Москва : Радио и связь, 1988. – 272 с.
6. Гостев, В. М. Система оптимизации проектирования сетей передачи данных / В. М. Гостев // Ученые записки Казанского университета. Серия: Физико-математические науки. – Казань, 2007. – №2. – С. 35 – 48.
7. Дубов, Ю. А. Многокритериальные модели формирования и выбора вариантов систем / Ю. А. Дубов, С. И. Травкин, В. Н. Якимец, – Москва : Наука, 1986. – 296 с.
8. Дуничкина, Н. А. Модели сравнения двух множеств оценок для бикритериальных задач синтеза оптимизированных по парето решений : текст / Н. А. Дуничкина, А. С. Куимова, – Нижний Новгород : Конгресс Международного форума «Великие реки», 2012. – 71 – 74 с.
9. Евтушенко, Ю. Г. Методы численного решения многокритериальных задач / Ю. Г. Евтушенко, М. А. Потапов // Доклады Академии наук СССР. – 1986. – Т. 291, № 1. – С. 25 – 39.
10. Зайченко, Ю. П. Структурная оптимизация сетей ЭВМ / Ю. П. Зайченко, Ю. В. Гонга. – Киев : Техника, 1986. – 168 с.
11. Канторович, Л. В. Применение математических методов в вопросах анализа грузоперевозок / Л. В. Канторович, М. К. Гавурин // Проблемы

повышения эффективности работы транспорта, Академия Наук СССР. – 1949. – С. 110 – 138.

12. Клейнрок, Л. Вычислительные системы с очередями / Л. Клейнрок – Москва : Мир, 1979. – 600 с.

13. Коган, Д. И. Динамическое программирование и дискретная многокритериальная оптимизация : учеб. пособие / Д. И. Коган. – Нижний Новгород : Изд-во ННГУ им. Н.И. Лобачевского, 2005. – 260 с.

14. Корбут, А. А. Дискретное программирование / А. А. Корбут, Ю. Ю. Финкельштейн; под. ред. Д. Б. Юдина. – Москва : Наука, 1969. – 368 с.

15. Кормен, Т. Алгоритмы: построение и анализ / Т. Кормен, Ч. Лейзерсон, Р. Ривест, К. Штайн // 2-е изд. – Москва : «Вильямс», 2006. – 1296 с.

16. Лазарев, Е. А. Алгоритм имитации отжига решения задачи оптимизации сети передачи данных / Е. А. Лазарев // Системы управления и информационные технологии. – 2012. – Т. 49, № 3. – 50 – 53 с.

17. Лазарев Е. А. Бикритериальная модель сети передачи данных / Е. А. Лазарев, Д. Е. Шапошников, П. В. Мисевич // Системы управления и информационные технологии. – 2011. – Т. 45, № 3.2. – 255 – 258 с.

18. Лазарев, Е. А. Генетические алгоритмы оптимизации сети передачи данных / Е. А. Лазарев, Д. Е. Шапошников, П. В. Мисевич // Системы управления и информационные технологии. – 2011. – Т. 46, № 4. – 59 – 63 с.

19. Лазарев, Е. А. Метод ветвей и границ для оптимизации структуры сети передачи данных / Е. А. Лазарев, Д. Е. Шапошников, П. В. Мисевич // Известия волгоградского государственного технического университета. – 2012. – Т. 10, № 14. – 189 – 193 с.

20. Лазарев, Е. А. Методы оценки эффективности алгоритмов решения многокритериальных задач / Е. А. Лазарев // Журнал Средневолжского математического общества. – 2012. – Т. 14, № 2. – 81 – 86 с.

21. Лазарев, Е. А. О применении эвристик для метода ветвей и границ решения задачи оптимизации сети передачи данных / Е.А. Лазарев // Труды

Нижегородского государственного технического университета им. Р.Е. Алексеева. – Нижний Новгород : НГТУ, 2012, №3. – 91 – 96 с.

22. Лазарев, Е. А. Применение метода ветвей и границ для оптимизации структуры сети передачи данных / Е. А. Лазарев, Д. Е. Шапошников, П. В. Мисевич // Информационные системы и технологии ИСТ-2012. Материалы XVIII международной научно-технической конференции. – Нижний Новгород : НГТУ, 2012. – С. 310 – 312.

23. Лопатин А. С. Метод отжига // Стохастическая оптимизация в информатике. – Санкт-Петербург : СПбГУ, Вып. 1, 2005. – 133 – 149 с.

24. Лотов, А. В. Многокритериальные задачи принятия решений : учеб. пособие / А. В. Лотов, И. И. Поспелова. – Москва : МАКС-Пресс, 2008. – 200 с.

25. Максименков, А. В. Основы проектирования информационно-вычислительных систем и сетей ЭВМ науч. изд. / А. В. Максименков, М. Л. Селезнев. – Москва : Радио и связь, 1991. – 320 с.

26. Ногин, В. Д. Принятие решений в многокритериальной среде: количественный подход / В. Д. Ногин. – 2-е изд., испр. и доп. – Москва : Физматлит, 2004. – 176 с.

27. Ногин, В. Д. Принятие решений при многих критериях / В. Д. Ногин. – Санкт-Петербург : Ютас, 2007. – 104 с.

28. Олифер, В. Г. Компьютерные сети. Принципы, технологии, протоколы / В. Г. Олифер, Н. А. Олифер. – Санкт-Петербург : Питер, 2006. – 864 с.

29. Подиновский, В. В. Парето-оптимальные решения многокритериальных задач / В. В. Подиновский, В. Д. Ногин. – 2-е изд., испр. и доп. – Москва : Физматлит, 2007. – 256 с.

30. Сигал, И. Х. Введение в прикладное дискретное программирование. Модели и вычислительные алгоритмы / И. Х. Сигал, А. П. Иванова. – Москва : Физматлит, 2007. – 304 с.

31. Стецко, А. А. Автоматизированное проектирование вычислительных сетей крупных проектных организаций / А. А. Стецко. – Ульяновск : УлГТУ, 2007. – 195 с.

32. Трей, Н. C# 2010: ускоренный курс для профессионалов : пер. с англ. / Н. Трей. – Москва : ООО "И.Д. Вильямс", 2010. – 592 с.
33. Хачатуров, В. Р. Комбинаторные методы и алгоритмы решения задач дискретной оптимизации большой размерности / В. Р. Хачатуров, В. Е. Веселовский, А. В. Злотов, С. У. Калдыбаев. – Москва : Наука, 2000. – 354 с.
34. Ху, Т. Целочисленное программирование и потоки в сетях / Т. Ху. – Москва : Мир, 1974. – 520 с.
35. Шефер, Е. А. Применение метода ветвей и границ для построения множества Парето в дискретной задаче векторной оптимизации / Е. А. Шефер. – Москва, 1981. – 9 с. - Деп. в ВИНТИ 10.01.81, № 402-81 Деп.
36. Штойер, Р. Многокритериальная оптимизация: теория, расчет и приложения : пер. с англ. / Р. Штойер. – Москва : Радио и связь, 1992. – 504 с.
37. Янбых, Г. Ф. Оптимизация информационно-вычислительных сетей / Г. Ф. Янбых, Б. А. Столяров. – Москва : Радио и связь, 1987. – 232 с.
38. Ingber, L. Simulated Annealing: Practice versus theory // L. Inberg, - USA : Mathematical and Computer Modelling, 18(11), 1993. – 29 – 57 p.
39. Ingber, L. Genetic Algorithms and Very Fast Simulated Reannealing: A Comparison // L. Inberg, B. Rosen, - USA : Mathematical and Computer Modelling, 16(11), 1992. – 87 – 100 p.
40. Orlin, J. B. Max flows in $O(nm)$ time or better / J. B. Orlin. – Palo Alto, CA, USA : Proceedings of the 2013 Symposium on the Theory of Computing, 2012. – 765 – 774 p.

ПРИЛОЖЕНИЕ А

Акт технической экспертизы



ВОСТОЧНО-СИБИРСКИЙ БАНК

660028, г. Красноярск, пр. Свободный, д. 46
Телефон: (391) 259-80-03,
Факс: (391) 259-82-17, 265-37-57,
E-mail: sberbank@vsb.sbrf.ru, www.sberbank.ru

В ИКИТ СФУ
г. Красноярск
ул. Академика Киренского 26Б

управление
отдел/сектор

УРС, Т ОСС, Т

№ *21-110X/183*

на №

от *22.06.2015*

АКТ ТЕХНИЧЕСКОЙ ЭКСПЕРТИЗЫ

О практическом применении полученных результатов
выпускной бакалаврской работы Тюрюминой М.Д.
на тему: «Оптимизация сети передачи данных на примере Восточно-Сибирского банка ОАО
Сбербанка России»

Настоящим подтверждаем, что результаты выпускной бакалаврской работы Тюрюминой М.Д. на тему: «Оптимизация сети передачи данных на примере Восточно-Сибирского банка ОАО Сбербанка России» обладают актуальностью, представляют практический интерес. Программная система, позволяющая оптимизировать сеть передачи данных разработанная Тюрюминой М.Д. может быть использована отделом систем связи и телекоммуникации с целью снижения затрат на каналы связи, повышения их производительности, что позволит осуществить существенную экономию и обеспечить требуемое качество. В результате применения разработанной программы были выявлены возможности совершенствования региональной сети.

Существенным достоинством является комплексность подхода к решению проблемы оптимизации и ее законченное решение

Директор управления технических
средств и телекоммуникаций



И.В. Лесков

ПРИЛОЖЕНИЕ Б

Входные данные для апробации работы программы

Таблица Б.1 – Данные об арендованных у провайдера «Ростелеком» каналах связи Сибирского банка на территории Республики Хакасии

Откуда	Куда	Пропускная способность канала, Кбит/с	Арендная плата в месяц, руб.
Абаза	Таштып	128	1417
Абакан	Абаза	2048	7848
Абакан	Белый Яр	2048	5400
Абакан	Очуры	1024	1000
Абакан	Подсинее	1024	2100
Абакан	Саяногорск	5120	14170
Абакан	Усть-Абакан	1024	1000
Абакан	Черемушки	2048	7000
Абакан	Черногорск	2048	2000
Абакан	Шира	2048	7848
Аскиз	Бирикчуль	128	5700
Аскиз	Бондарево	512	7085
Белый Яр	Аскиз	256	5756
Белый Яр	Боград	2048	7085
Бискамжа	Сорск	1024	6993
Саяногорск	Бея	1024	5995
Саяногорск	Майна	1024	5995
Черемушки	Бельтирское	2048	3200
Черемушки	Бискамжа	1024	4300
Черемушки	Тёя	512	2779

ПРИЛОЖЕНИЕ В

Географическое положение узлов коммутации и каналов связи на территории Республики Хакасии

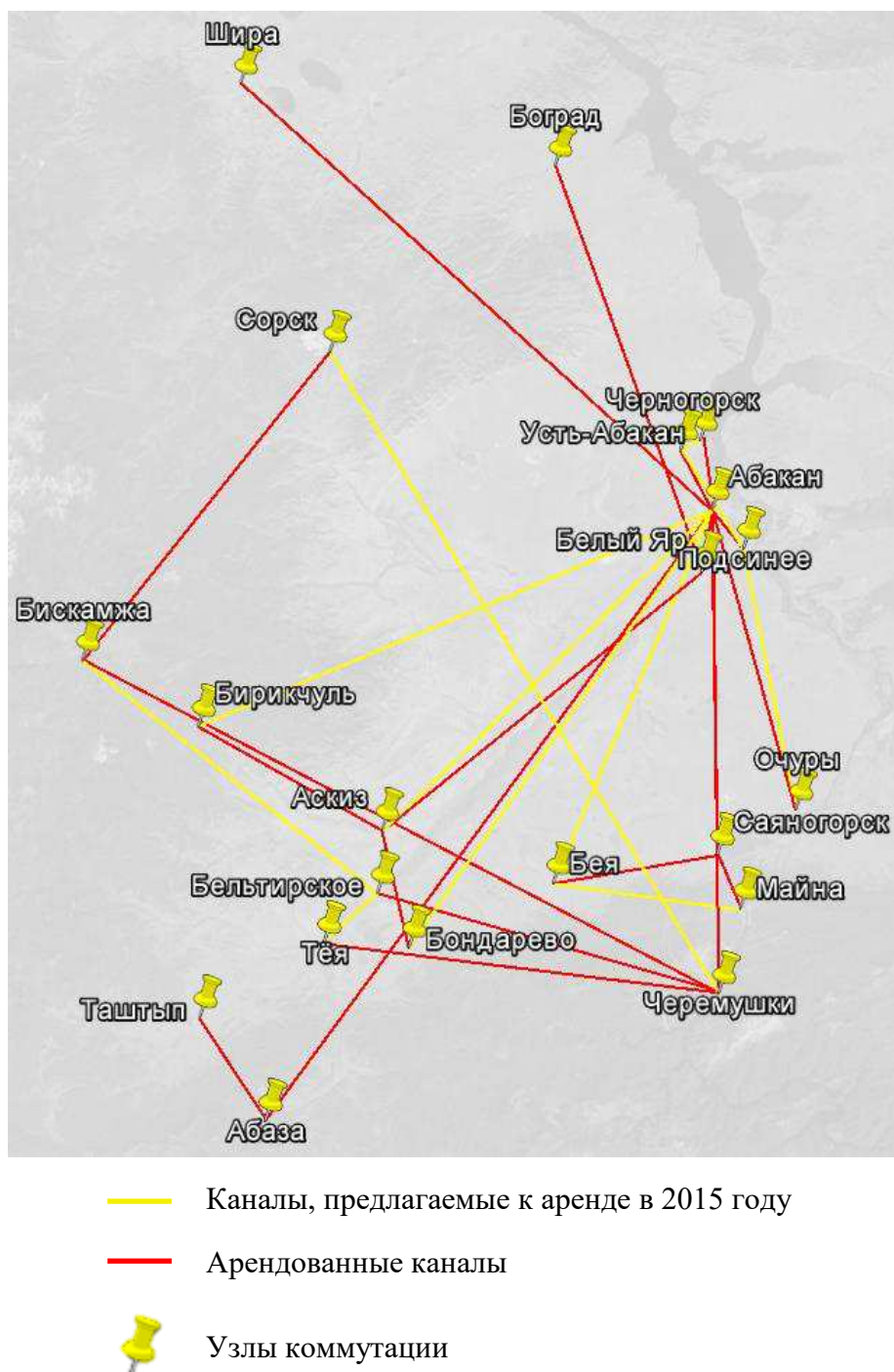


Рисунок В.1 – Расположение узлов коммутации, арендованных и предлагаемых к аренде каналов связи на территории Республики Хакасии

Федеральное государственное автономное
образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»
Институт космических и информационных технологий
Кафедра «Информатика»

УТВЕРЖДАЮ

Заведующий кафедрой

А.С. Кузнецов

инициалы, фамилия

подпись

«16»

06

2017 г.

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Оптимизация сети передачи данных на примере Сибирского банка ПАО

Сбербанка России

27.04.03 Системный анализ и управление

27.04.03.02 Системный анализ данных и технологий принятия решений

Научный руководитель

А.А. Даничев

16.06.17

доцент, к.т.н.

подпись, дата

должность, ученая степень

А.А. Даничев

инициалы, фамилия

Выпускник

М.Д. Тюрюмина

16.06.17

подпись, дата

М.Д. Тюрюмина

инициалы, фамилия

Рецензент

Д.О. Видинеев

15.06.17

вед. инж. ЦИ ПАО Сбербанк

подпись, дата

должность, ученая степень

Д.О. Видинеев

инициалы, фамилия

Красноярск 2017